DOCUMENT RESUME

ED 175 577 PS 010 898

AUTHOR Lazar, Irving; And Others
TITLE Lasting Effects After Preschool. A Report of the Consortium for Longitudinal Studies.
INSTITUTION Education Commission of the States, Denver, Colo.
SPONS AGENCY Administration for Children, Youth, and Families (DHEW), Washington, D.C.; New Jersey State Dept. of Higher Education, Trenton.; William and Flora Hewlett Foundation, Palo Alto, Calif.
REPORT NO OHDS-79-30178
PUB DATE Oct 78
NOTE 337p.; For summary report, see PS 010 386; For related documents, see ED 148 471-72
AVAILABLE FROM Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 (Stock Number 017-090-00047-0)

EDRS PRICE MF01/PC14 Plus Postage.
DESCRIPTORS *Academic Achievement; Achievement Need; Achievement Rating; Disadvantaged Youth; Early Childhood Education; Followup Studies; Grade Repetition; Intelligence Quotient; *Intervention; *Longitudinal Studies; *Low Income Groups; *Preschool Programs; *Program Effectiveness; Special Education

ABSTRACT

This second general technical report of the Consortium for Longitudinal Studies summarizes the findings of current analyses of longitudinal studies of low income children who participated in experimental preschool programs initiated in the 1960s and includes additional data and further analyses of the 1976-77 follow-up study. The pooled data bank of the 12 research groups making up the Consortium provided information on approximately 3,000 low income children who either participated in early intervention programs or served as controls. The common information across projects includes various pre-enrollment measures of the children's home background, a pretest IQ score, and at least one IQ test score collected immediately after the preschool experience. Each investigator also collected a variety of cognitive and behavioral measures which are not common across all projects. In 1976-77, members of the Consortium collaborated in a common follow-up data collection effort. These data include both child and parent interviews, children's IQ and achievement test scores, and information on whether the children had ever failed a grade or been assigned to special education classes. The children were aged 9 through 19 years old at the 1976-77 follow-up. The findings showed that high quality early education programs for low income children had lasting effects in five areas: (1) reducing the number of children assigned to special education classes, (2) reducing the number of children retained in grade, (3) increasing children's math achievement scores at fourth grade, (4) increasing IQ scores at least up to age 13, and (5) influencing aspects of children's and mothers' achievement orientation. (Author/JMB)

ED175577

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Lasting Effects After Preschool

A report of the Consortium for
Longitudinal Studies* under the
supervision of Irving Lazar
and Richard B. Darlington
October, 1978

Further analyses of
longitudinal studies
conducted by:

PHYLLIS LEVENSTEIN
LOUISE MILLER
FRANCIS PALMER
DAVID WEIKART
MYRON WOOLMAN
EDWARD ZIGLER
KUNO BELLER
CYNTHIA DEUTSCH
MARTIN DEUTSCH
IRA GORDON
SUSAN GRAY
MERLE KARNES

A project of the Education Commission of the States, this research is supported in part by grants from the Administration for Children, Youth and Families (OHDS; DHEW) and the William and Flora Hewlett Foundation and the New York State College of Human Ecology at Cornell University.

* *Formerly the Consortium on Developmental Continuity.*

PS010898

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Office of Human Development Services
Administration for Children, Youth and Families

DHEW Publication No. (OHDS) 79-30178

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY
Education Commission
of the States
TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Members of the Consortium for Longitudinal Studies
(formerly, the Consortium on Developmental Continuity)

Kuno Beller
Psychology Department
Temple University
Philadelphia, PA 19122

Richard B. Darlington
Psychology Department
Cornell University
Ithaca, NY 14853

Martin & Cynthia Deutsch
School of Education
New York University
239 Green St.
New York, NY 10003

Ira Gordon*
University of North Carolina
Chapel Hill, NC 27514

Susan Gray
George Peabody College
Box 151
Nashville, TN 37203

Merle Karnes
Colonel Wolfe Preschool
403 E. Healey
Champaign, IL 61820

Irving Lazar
N-135 MVR Hall
Cornell University
Ithaca, NY 14853

Phyllis Levenstein
Verbal Interaction Project
5 Broadway
Freeport, NY 11520

Louise Miller
Psychology Department
University of Louisville
Louisville, KY 40208

Francis Palmer
Merrill-Palmer Institute
71 E. Ferry
Detroit, MI 48202

David Weikart
High/Scope Foundation
600 N. River
Ypsilanti, MI 48197

Myron Woolman
Institute for Educational Research
4828 16th St.
Washington, DC 20011

* Deceased, September, 1978. Address inquiries to:

Robert Emile Jester
417 N.E. 4 Ave.
Gainesville, FL 32601

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402

Stock Number 017-000-00047-0

Professional Staff of the Consortium

Central Staff (located at Cornell University)

Irving Lazar, Ph.D. and Richard B. Darlington, Ph.D., Co-Directors

Daniel Koretz, Ph.D.
Harry Murray
Donald Poe
Marilyn Rosche
Jacqueline Royce, Ph.D.
Ann Snipper, Ph.D.
Patrick Vitale
Judy Vopava, Ph.D.
Agelia Ypelaar

Support Staff: Ann Bell, Nancy Burston, Jane Pedersen, Marjorie Wikerd, and Chris Wojno

Staff for Individual Projects

Marcene Root (Beller)
Bonnie Fattell (Beller)
John McNichol (Beller)
Ronnie Greenstein (Deutsch)
Deidre Hamlin (Deutsch)
Judy MacMurray (Gordon)
Pat Schlenker (Gordon)
Barbara Ramsey, Ph.D. (Gray)
William Hallau (Karnes)
Charles Plummer, Ph.D. (Karnes)
John Madden, Ph.D. (Levenstein)
Kay Proctor (Miller)
Roni Bizzell (Miller)
Bobbie Jones, Ph.D. (Miller)
Georgia Brandstadter, Ph.D. (Palmer)
Ronald Siegel (Palmer)
Thelma Semlear, Ph.D. (Palmer)
Lawrence Schweinhart, Ph.D. (Weikart)
Burness Broussard, Ph.D. (Woolman)
Victoria Seitz, Ph.D. (Zigler)
Penelope Trickett, Ph.D. (Zigler)

IN MEMORY

This report is dedicated to our friend and colleague, Dr. Ira Gordon, who died suddenly of a heart attack on September 7, 1978. Ira's contributions to the well being of children and his pioneering work in improving the lot of the poor will remain classic chapters in the history of child development. His wisdom, compassion and good sense played a significant role in the success of this collaborative effort. We miss him sorely.

ACKNOWLEDGEMENT

The solutions to the complex problems of secondary analyses posed in this study and the extraordinary level of recovery of the original subjects were accomplished through the help of a great many people.

The individual investigators, their staff and students worked under enormous handicaps which they largely surmounted.

Dr. James Peterson and Dr. Homer Elseroad of the Education Commision of the States played key roles in supporting cooperation of the hundreds of public schools who participated in this study. Dr. Edith Grotberg, Dr. Ray Rackley and Dr. Bernard Brown of the Administration for Children, Youth and Families not only made possible the financial underpinning of this work, but provided professional counsel and consultation throughout our work. Dr. Brown particularly gave generously of his own time and methodological skills as a critic and consultant at different points in the treatment of data.

A number of distinguished professionals have served as consultants, critics, and respondents at professional meetings and have given of their time to do careful, independent study of various parts of our work. We wish particularly to acknowledge the contributions in this regard of Dr. Robert McCall of Boys Town; Dr. Shirley Moore, University of Minnesota; and Dr. Virginia Shipman and Dr. Irving Sigel of the Educational Testing Service.

A number of our colleagues at Cornell have also provided advice, consultation, and criticism throughout our work. Their willingness to be immediately available, to be tough-minded and constructive volunteers has helped strengthen the scientific value of this work in ways that are incalculable. Included are Professors Urie Bronfenbrenner, Steven Caldwell, John Doris, Jeanne Mueller, and Henry Ricciuti.

PREFACE

This is the second general technical report of the Consortium for Longitudinal Studies. The first report entitled "The Persistence of Preschool Effects" was published in October of 1977. While the main findings are summarized in this volume, the reader wishing more technical information on the earlier findings should refer to that report. This volume includes additional data and further analyses of the 1976-77 follow-up study conducted by members of the Consortium.

Lasting Effects After Preschool

1977-78 Report

Abstract

I. Overview

The Consortium for Longitudinal Studies is a collaborative effort of twelve research groups conducting longitudinal studies on the outcomes of early education programs. This report summarizes the findings of current analyses of longitudinal studies of low income children who participated in experimental infant and preschool programs initiated in the 1960's.

II. Methods

The central data bank contains information on approximately 3,000 low-income children. These children participated in early intervention programs or served as controls. The common information across projects includes various measures of the children's home backgrounds collected before enrollment in preschool, for most cases a pretest IQ score, and at least one IQ test test score collected immediately after the preschool experience. In addition, each investigator collected a variety of cognitive and behavioral measures which are not common across all projects. In 1976-77, members of the Consortium collaborated in a common follow-up data collection effort. These data include interviews with both children and their parents, WISC-R IQ test scores for the children, achievement test scores from the schools, and indicators of the children's standing in their schools (i.e., whether they had ever failed a grade or been assigned

to special education classes). The children were aged 9 through 19 years old at the 1976-77 follow-up.

Methodological problems such as problems of sample selection and measurement are discussed. Very detailed attrition analyses are presented. The statistical techniques used to test hypotheses and pool results are detailed. For example, we never pooled all subjects together. Instead, each project was considered separately for each hypothesis test. Then the p values of each project were pooled in order to test the null hypothesis that there is no average effect of preschool across programs.

III. Findings

The detailed attrition analyses indicated that attrition was essentially random, introducing no noticeable biases into our other analyses.

The data analyzed thus far show that early education programs for children had lasting effects in the following areas:

1. Assignment to special education. Early education programs significantly reduced the number of children assigned to special education classes. This result was true after controlling for the effects of children's initial IQ scores, sex, ethnic background and family background. It held even after controlling for children's IQ scores at age 6. Furthermore, the benefit apparently extended to all the low-income participants, regardless of their initial abilities or early home backgrounds.

2. Retention in grade (grade failure). The combined evidence from eight projects able to collect this information indicates that early education significantly reduced the number of children retained in grade. Again, the result was true when measures of early child characteristics and home background variables were controlled. Furthermore, all low-income children -- regardless of sex, ethnic background, early IQ, and home background -- benefitted in this way.

3. Achievement test scores. The Consortium had the most information for children at the fourth grade level. The combined evidence from projects able to collect this information indicates that early education significantly increased children's scores on fourth grade mathematics achievement tests with a suggestive trend toward increased scores on fourth grade reading tests.

4. Intelligence test scores. Low-income children who attended preschools surpassed their controls on the Stanford-Binet IQ test for up to 3 years after the programs ended. Current Wechsler IQ scores show that the children maintained that superiority in the Gordon, Levenstein, and Palmer projects. There were no treatment/control differences found in projects whose subjects were aged 13 or older. Using WISC scores as outcomes, there was no evidence that preschool benefitted boys more than girls, or vice-versa; or that children whose mothers had different levels of education were helped differentially.

5. Non-cognitive measures. Children who attended preschool were more likely than control children to give achievement-type reasons for being proud of themselves. The family context also appears to have been affected. Specifically, mothers of children who attended preschool had higher vocational aspirations for their children than the children had for themselves. This discrepancy was not found in mothers of control children.

6. Use of child welfare services. For the four projects with data available, a preliminary investigation found no significant treatment/control differences in families' use of Title IV child welfare services.

In order to illuminate the means by which preschool exerts its impact, ten different characteristics of preschool programs were examined: age of entry, length of program (in years, months per year, and hours per year), degree of parental influence, location of program, professional vs. paraprofessional staff, preservice training of staff, language goals for children, and amount of teaching structure. None of these variables emerged as more effective than the others when assignment to regular vs. special education classrooms was the criterion of effectiveness. Furthermore, no one type of program was more effective (using the same criterion) than another with certain kinds of children (i.e. children differing on initial IQ scores, sex, and family background measures). We concluded that these high-quality programs were apparently about equally effective in helping low-income children.

IV. Implications

High quality early education programs are likely to benefit both low-income children and the larger society by: reducing the number of children in later costly special education programs in schools, helping children avoid grade failure, increasing children's math achievement scores at fourth grade and IQ scores at least up to age 13, and influencing aspects of children's and mothers' achievement orientation.

TABLE OF CONTENTS

Presentations and Publications by Consortium Members and Staff, August, 1977 - August, 1978 i

List of Tables and Figures vi

I. INTRODUCTION 1

II. DESCRIPTION OF THE CONSTITUENT STUDIES 5

III. METHODS 28

- A. General Problems in Longitudinal Assessments of Program Effects 28
- B. The Problem of Attrition 30
- C. Problems of Measurement 30
- D. Issues Inherent in Multi-sample Secondary Analysis 32

IV. COGNITIVE EFFECTS OF PRESCHOOL 37

- A. Intelligence Test Scores 37
- B. Assignment to Special Education and Retention in Grade 53
- C. Achievement Tests 75

V. NON-COGNITIVE EFFECTS OF PRESCHOOL 82

- A. Attitudes and Values 82
- B. Use of Title IV Child Welfare Services 116

VI. DIFFERENTIAL EFFECTS OF PRESCHOOL 122

- A. Which Children Benefit Most on WISC-R Test Scores? 122
- B. Which Children Benefit Most on School Performance? 131
- C. What Kinds of Programs Were Most Effective? 138
- D. Matching Children to Programs 150
- E. Summary of Differential Effects 156

VII. THE EFFECTS OF FAMILY CHARACTERISTICS ON SCHOOL OUTCOMES FOR LOW-INCOME CHILDREN 158

- A. Predicting School Outcomes 158

VIII. REVIEW OF THE DETERMINANTS OF SPECIAL EDUCATION PLACEMENT 165

IX. SUMMARY AND IMPLICATIONS 171

- A. Implications 178

REFERENCES 182

APPENDIX A: Attrition 189

APPENDIX B: Status Report by Instrument 221

APPENDIX C: Response Rates and Final Dispositions 223

APPENDIX D: Findings From the Deutsch Project 227

APPENDIX E: Interview Coding and Analysis of Construct Validity 230

APPENDIX F: Instruments 252

Presentations and Publications by Consortium Members and Staff

August, 1977 - August, 1978

Presentations by Consortium members were given at the meetings of the American Psychological Association (August, 1977 and August, 1978); the American Association for the Advancement of Science (February, 1978); the Colorado Association for the Education of Young Children (February, 1978); the American Educational Research Association (March, 1978); the Minnesota Round Table on Early Childhood Education (May, 1978); and the American Statistical Association (August, 1978). In addition, conferences were held with staffs of the Governors' offices of Wyoming, Colorado, and Nebraska (August, 1978) and of North Carolina, South Carolina, Georgia, Tennessee, Kentucky, and Virginia (May and June, 1978). Additional presentations were given at a Conference on Parenting Education held at Wheelock College, Boston, Massachusetts (March, 1978) and at the NICHD Conference on Prevention of Retarded Development in Psychosocially Disadvantaged Children at the University of Wisconsin, Madison (July, 1978).

Available papers and publications are listed below:

Brown, B. Long-term gains from early intervention: An overview of current research. In B. Brown (Ed.), Found: Long-term gains from early intervention. Boulder, Colorado: Westview Press, 1978.

Consortium on Developmental Continuity. The persistence of preschool effects. Final Report, Grant No. 18-76-07843 to the Administration for Children, Youth and Families, Office of Human Development Services, U.S. Department of Health, Education and Welfare, September, 1977.

Darlington, R. B. Member of the Interagency Panel on Early Childhood Education. June, 1978.

Darlington, R.B. Methods, issues, and some illustrative findings in analyzing the data of the Consortium on Developmental Continuity. Presented at the annual meeting of the American Educational Research Association, Toronto, March, 1978(a).

Darlington, R.B. Multiple comparisons and related problems. Presented at the annual meeting of the American Psychological Association, Toronto, August, 1978(b).

Gordon, I.J. Early intervention: Matching programs to children. Presented at the annual meeting of the American Association for the Advancement of Science, Washington, D.C., February, 1978.

Gordon, I.J., & Guinagh, B.J. School performance as a function of early stimulation. Presented at the annual meeting of the American Psychological Association, San Francisco, August, 1977.

Gordon, I.J., & Guinagh, B.J. *Middle school performance as a function of infant stimulation*. Final report to the Administration for Children, Youth and Families, Department of Health, Education and Welfare, Project No. NIH-HEW-OCD-90-C-908, March 31, 1978.

Gray, S.W., Klaus, R.A., & Ramsey, B.K. Participants in The Early Training Project: 1962-1977. Presented at The NICHD Conference on Prevention of Retarded Development in Psychosocially Disadvantaged Children, University of Wisconsin, Madison, March, 1978.

Gray, S.W., Klaus, R.A., & Ramsey, B.K. The Early Training Project participants at the end of high school. In Kuno Beller (Chair), "Long-term Gains in Early Intervention", symposium presented at the annual meeting of the American Educational Research Association, Toronto, March, 1978.

Hubbell, V.R. The Developmental Continuity Consortium Study: Secondary analysis of early intervention data. In B. Brown (Ed.), *Found: Long-term gains from early intervention*. Boulder, Colorado: Westview Press, 1978, pp. 111-128.

Hubbell, V.R. Different types of early intervention programs: How children and parents react to them. *Children Today*, in press.

Hubbell, V.R. Differential effects of three early childhood intervention delivery systems. Paper presented at the Annual Meetings of the American Association for the Advancement of Science, Washington, D.C., February, 1978.

Hubbell, V.R. Early intervention research and family impact. Paper presented at the Groves Conference on Marriage and the Family, Washington, D.C., April 27-28, 1978.

Karnes, M.B., Plummer, C.M., & Lee, R.C. Immediate, short-term and long-range effects of five preschool programs for disadvantaged children. In Kuno Beller (Chair), "Long-term Gains in Early Intervention," symposium presented at the annual meeting of the American Educational Research Association, Toronto, March, 1978.

Koretz, D., Vopava, J., & Darlington, R. Matching early education programs to children: Some methodological limitations. Presented at the annual meeting of the American Association for the Advancement of Science, Washington, D.C., February, 1978.

Lazar, I. Participant in the meeting of the Interagency Panel on Early Childhood Education, June, 1978.

Lazar, I., Koretz, D., Royce, J., & Beller, E.K. "Socialization for school success: The long-term impact of early intervention on values, motivation, and actual school performance. In D. McClelland (Ed.), Education for Values, in press.

Levenstein, P. Third grade effects of the Mother-Child Home Program. Final Report to the Developmental Continuity Consortium, 1978.

Miller, L.B. Differences in process and product in a four-program comparison: The Louisville Study. Paper presented at the Annual Meetings of the American Educational Research Association, New York, April, 1977.

Palmer, F.H. The effects of minimal early intervention on subsequent school performance and IQ. Final Report Contract No. 26-77-09617, Educational Commission of the States. New York: Palmer Associates, 1977.

Palmer, F.H. "The effects of early childhood intervention." In B. Brown (Ed.), Found: Long-term gains from early intervention. Boulder, Colorado: Westview Press, 1978, pp. 11-35.

Palmer, F.H. Early intervention treatments which have been tried, documented, and assessed. Background paper presented at the NICHD Conference on Prevention of Retarded Development in Psychosocially Disadvantaged Children, University of Wisconsin, Madison, July, 1978.

Palmer, F.H. & Anderson, L.W. The effectiveness of early childhood intervention: Findings from longitudinal studies. In E. Zigler & J. Valentine (Eds.), *Project Head Start: A legacy of the war on poverty*. N.Y.: Free Press, in press.

Palmer, F.H. & Siegel, R.J. Harlem Study follow up: Cognitive and social characteristics of the sample at ages 13 to 15. Final Report, Grant No. 90-C1274(01) from the Administration for Children, Youth and Families (OHDS, DHEW), October, 1978.

Schweinhart, L.J. The impact of the evaluation of the Ypsilanti Perry Preschool Project. In S.A. Raizen (Chair), "The impact of evaluation: Lessons drawn from the evaluations of five early childhood education programs," symposium presented at the annual meeting of the American Educational Research Association, Toronto, March, 1978.

Schweinhart, L.J. & Weikart, D. The lasting scholastic impact of preschool. Paper presented at the Annual Meetings of the American Educational Research Association, Toronto, 1978.

Schweinhart, L.J. & Weikart, D. Proyecto Pre-escolar Ypsilanti-Perry: ¿Puede la educación pre-escolar producir una diferencia duradera en al futuro educativo del niño: *Educación Hoy: Perspectivas Latino-americanas*, Otoño 7, Na. 42, Nov.-Dic. 1977, pp. 32-42.

Schweinhart, L.J., & Weikart, D.P. Perry Preschool effects nine years later: What do they mean? Presented at the NICHD Conference on Prevention of Retarded Development in Psychosocially Disadvantaged Children, University of Wisconsin, Madison, July, 1978.

Seitz, V., Apfel, N., & Efron, C. Long-term effects of early *intervention: The New Haven Project*. In B. Brown (Ed.), *Found*: Long-term gains from early intervention. Boulder, Colorado: Westview Press, 1978, pp. 79-109.

Seitz, V., Apfel, N., & Rosenbaum, L.K. Projects Head Start and Follow Through: A longitudinal evaluation of adolescents presented at the NICHD Conference on Prevention of Retarded Development in Psychosocially Disadvantaged Children, University of Wisconsin, Madison, July, 1978.

Vopava, J., & Royce, J. Comparison of the long-term effects of infant and preschool programs on academic performance. Presented at the annual meeting of the American Educational Research Association, Toronto, March, 1978.

Weber, C.V., Foster, P.W., & Weikart, D.P. An economic analysis of the Ypsilanti Perry Preschool Project. Monographs of the High/Scope Educational Research Foundation, No. 5, 1978.

Weikart, D.P., Bond, J.T., & McNeil, J.T. The Ypsilanti Perry Preschool Project: Preschool years and longitudinal results through fourth grade. Monographs of the High/Scope Educational Research Foundation, No. 3, 1978.

Weikart, D., Deloria, D., & Lawser, S. "Longitudinal results of the Ypsilanti Perry Preschool Project." Monographs of the High/Scope Educational Research Foundation, No. 1. Ypsilanti.

Weikart, D.P., Epstein, A.S., Schweinhart, L. J., & Bond, J.T. The Ypsilanti Preschool Curriculum Demonstration Project: Preschool years and longitudinal results. Monographs of the High/Scope Educational Research Foundation, No. 4, 1978.

Ypelaar, A. Using the jacknife in an analysis of preschool program characteristics. Presented at the annual meeting of the American Statistical Association, San Diego, August, 1978. (To appear in the 1978 proceedings of the Social Statistics Section - American Statistical Association.)

Zigler, E., & Trickett, P.K. IQ, social competence, and evaluation of early childhood education programs. American Psychologist, 1978, 33, 789-798.

List of Tables and Figures

Table 1: Characteristics of Early Education Programs and Ages of Subjects for Each Data Set 25

Table 2: Background Characteristics, IQ Scores, and Preschool Attendance for Each Data Set 27

Table 3: The Effect of Early Education on IQ Score at Age 6 When Background Variables and Pretest IQ Are Controlled 45

Table 4: The Effect of Early Education on IQ Score at Age 6 When Background Variables Alone Are Controlled 46

Table 5: Mean WISC IQ Scores, Program vs. Control by Project 47

Table 6: Mean WISC IQ Scores, Program vs. Control by Project 48

Table 7: Standard Deviations of WISC IQ Scores, Program vs. Control by Project 50

Table 8: Standard Deviations of WISC Subtest Scores, Program vs. Control by Project 51

Table 9: Percentage of Subjects Placed in Special Education Classes, Treatment vs. Control 69

Table 10: Percentage of Subjects Retained in Grade, Treatment vs. Control 70

Table 11: Percentage of Underachieving Students (Placed in Special Education Classes and/or Retained in Grade and/or Dropped Out of School), Treatment vs. Control 71

Table 12: The Effect of Early Education on School Outcomes When Background Variables and Pretest IQ Score Are Controlled 72

Table 13: The Effect of Early Education on School Outcomes When Background Variables Alone Are Controlled 73

Table 14: The Effect of Early Education on School Outcomes When Background Variables and IQ at Age 6 Are Controlled 74

Table 15: Fourth Grade Treatment-Control Differences on Achievement Tests 81

Table 16: Dependent Variables for Analysis on Non-Cognitive Outcomes 99

Table 17: Educational and Occupational Attainment of Parents and Aspirations of Children, Treatment and Control Pooled 101

Table 18: Summary of Treatment Effects on Family Context of Achievement Orientation 102

Table 18A: Mean of Mother's Occupational Aspirations for Child, Treatment vs. Control 103

Table 18B: Mean Discrepancy Between Mother's Occupational Aspirations for Child and Child's Own Aspirations, Treatment vs. Control 104

Table 19: Summary of Treatment Effects on Achievement-Orientation Variables 105

Table 19A: Mean of Children's Occupational Aspirations, Treatment vs. Control 106

Table 19B: Percent of Children Giving Achievement-Related Reasons for Being Proud of Themselves, Treatment vs. Control 107

Table 19C: Percent of Children Giving Achievement-Related Reasons for Being Proud of Themselves, by Sex, Age, and Treatment/Control Status 108

Table 20: Mean of Children's Self-Evaluations of Their Overall School Performance Relative to their Peers, Treatment vs. Control 109

Table 21: Summary of Treatment Effects on Self-Evaluation and Sociability Variables 111

Table 21A: Percent of Respondents Who Stated that They Have "Special Friends" With Whom they Spend Time, Treatment vs. Control 112

Table 21B: Mean Frequency with which Respondents Spend Time with "Special Friends" (Times per month), Treatment vs. Control 113

Table 21C: Mean Sociability Ratings of "The Best Things about School," Treatment vs. Control 114

Table 21D: Mean Sociability Ratings of Spare-Time Activities, Treatment vs. Control 115

Table 22: Correlation of Sex with WISC Subtest Scores, by Project 127

Table 23: Correlation of Sex with WISC Subtest Scores, by Project 128

Table 24: Correlation of Mother's Education with WISC IQ Scores, by Project 129

Table 25: Correlation of Mother's Education with WISC Subtest Scores, by Project 130

Table 26: Relationship Between School Outcomes and the Set of Treatment by Background Crossproducts (Interactions) When Linear Terms are Partialled Out 135

Table 27: The Effect of Initial IQ Level on the Relationship Between Early Education and School Outcomes 136

Table 27A: F Values for Beta Weight of Interaction Terms (Crossproducts of Treatment and Background Characteristics 137

Table 28: Subgroups and Effectiveness Scores 147

Table 29: Reduced Set of Program Variables Included in the Analysis 148

Table 30: Jacknifed Estimates of the Effectiveness Correlation and Associated Probability Level for Each Program Variable 149

Table 31: *F* Values and *p* Values for Incremental R^2 of Interaction Terms Using Special Ed as Dependent Variable 154

Table 32: *F* Values and *p* Values for Incremental R^2 of Interaction Terms Using Special Ed as Dependent Variable 155

Table 33: The Effect of Mother's Education on School Outcomes When Preschool Attendance and Other Background Variables Are Controlled 161

Table 34: The Effect of Mother's Education on School Outcomes When Preschool Attendance, Other Background Variables, and IQ at Age 6 Are Controlled 162

Table 35: The Effect of Other Background Variables on School Outcomes When Mother's Education and Preschool Attendance Are Controlled 163

Figure 1: Assignment to Special Education Classes: Diagram Showing Network of Variables Suggested by Data, Significant Paths Only 165

Table A-1: Final Samples as a Percent of Original Sample by Project 201

Table A-2: Final Sample as a Percent of Original Sample for Interview, by Project 202

Table A-3: Final Sample as a Percent of Original Sample for Youth Interview, by Project 203

Table A-4: Final Sample as a Percent of Original Sample for School Record Form, by Project 204

Table A-5: Final Sample as a Percent of Original for WISC-R by Project 205

Table A-6: General Attrition - Comparison of Mother's Education 206

Table A-7: General Attrition - Comparison of Hollingshead ISP Score 207

Table A-8: General Attrition - Comparison of Pretest Stanford-Binet IQ Scores 208

Table A-9: Attrition on the Parent Interview - Comparison of Mother's Education 209

Table A-10: Attrition on the Parent Interview - Comparison of Hollingshead ISP Scores 210

Table A-11: Attrition on the Parent Interview - Comparison of Pretest Stanford-Binet Scores 211

Table A-12: Attrition on the Youth Interview - Comparison of Mother's Education 212

Table A-13: Attrition on the Youth Interview - Comparison of Hollingshead ISP Scores 213

Table A-14: Attrition on the Youth Interview - Comparison of Pretest Stanford-Binet IQ Scores 214

Table A-15: Attrition on the School Record - Comparison of Mother's Education 215

Table A-16: Attrition on the School Record - Comparison of Hollingshead ISP Scores 216

Table A-17: Attrition on the School Record - Comparison of Pretest Stanford-Binet IQ Score 217

Table A-18: Attrition on the WISC-R - Comparison of Mother's Education 218

Table A-19: Attrition on the WISC-R - Comparison of Hollingshead ISP Scores 219

Table A-20: Attrition on the WISC-R - Comparison of Pretest Stanford-Binet IQ Scores 220

Table B-1: Number of Cases Received by September 30, 1978 for Each Instrument by Project 222

Table C-1: Response Rate Computations and Final Dispositions for Eight Sites 226a

Table E-1 Percent of School "Achievers" and "Non-Achievers" Who Gave Achievement-Related Reasons for Being Proud of Themselves 247

Table E-2 Mean of Children's Self-Evaluations of Their Overall School Performance Relative to Their Peers, "Achievers" vs. "Non-Achievers" 248

Table E-3 Mean Achievement-Orientation Ratings of Spare-Time Activities, "Achievers" vs. "Non-Achievers" (With competitive athletics coded as achievement-related) 249

Table E-4 Mean Achievement-Orientation Ratings of Spare-Time Activities, "Achievers" vs. "Non-Achievers" (Omitting competitive athletics from the achievement-related category) 250

Table E-5 Mean Discrepancy Between Mother's Occupational Aspirations for the Child and the Child's Own Aspirations, "Achievers" vs. "Non-achievers" (Based on Hollingshead Codes) 251

INTRODUCTION

In the early 1960's, a mood of exuberant optimism was in the air regarding society's ability to solve problems. Technological know-how had thrust American men into orbit less than 5 years after the launching of Sputnik I. President Kennedy had successfully applied Keynes' theories of economic growth to propel the nation out of recession. Thus, President Johnson's announcement of the War on Poverty in 1964 was met with hope and confidence that the know-how of social scientists -- together with infusions of dollars -- would substantially alleviate if not eliminate the causes of poverty. As part of that effort, Project Head Start was initiated in 1965 as a compensatory educational program for low-income preschoolers and their families. Its aims included the stimulation of the children's social and cognitive development, the provision of health services, and the encouragement of parental involvement with their children and in the community.

One year after the first man walked on the moon, the Westinghouse/Ohio evaluation report of effectiveness of Head Start programs appeared (Cicirelli, _et_ _al_., 1969). Essentially, it concluded that Head Start programs had only a few weak and fleeting effects.[1] Shortly thereafter, coinciding with efforts by a new administration to dismantle many anti-poverty programs, Moynihan advised a national policy of "benign neglect" of poor Blacks; Jencks, et al. (1972) suggested that indices of school quality were unrelated to children's level of academic achievement; and Jensen (1969) argued that intelligence is primarily genetically determined.

---

[1] While these findings have been extensively criticized by methodologists and others, many have overlooked the report's influence in defining the criteria of "success": an increase in cognitive development which is maintained long after children leave the program (Datta, 1976).

Since that time, slowly, but perceptibly, the climate of opinion has begun to improve. While no longer exhibiting the unbridled optimism of the 1960's, a number of social scientists have sounded cautious notes to the effect that early intervention programs may provide benefits which endure over time (Brown, 1978; Datta, n.d.; Mann, et al., 1976). The research reported by the Consortium on Developmental Continuity (1977) contributed significantly to this assessment. The Consortium[1] consists of 12 different investigators who independently conducted experimental preschool intervention programs in the early and mid-1960's. Two additional members provide coordination and data analysis.

In 1975 the Consortium members agreed to pool their original data and to collect common follow-up data in 1976-77. The initial findings of the follow-up effort indicated that preschool intervention programs had significant long-term effects on school performance. Specifically, compared to children in control groups, low-income children who received early education were better able to meet the minimal requirements of their schools, as shown in reduced rates of assignment to special education classes and reduced in-grade retention. In addition, the gains on IQ scores achieved by children in preschool programs were maintained for at least three years after leaving the program (Consortium, 1977).

---

[1] The name of the Consortium has recently been changed to the Consortium for Longitudinal Studies.

This report describes further analyses and findings of the Consortium's 1976-77 follow-up data. Work over the past year has concentrated on three general areas: (a) re-analysis of questions treated in the earlier report (Consortium, 1977) in order to include raw data received after July, 1977; (b) a search for long-term measurable effects of preschool intervention programs beyond those considered in the earlier report; and (c) a search for evidence of differential effectiveness of preschool intervention programs. In preview we may say that our earlier findings have been confirmed. In addition, there is evidence that preschool programs affected children's achievement test scores, children's achievement orientation, and parents' aspirations for their children's future vocations. Using rate of assignment to special education classes and grade failure as measures of effectiveness, there was no evidence that programs were differentially effective for children from different home backgrounds.

In this report we describe in detail each of the analyses and their limitations so that technical readers can assess our methods and conclusions for themselves. In addition, readers are asked to keep a few caveats in mind as they read this report.

1. While the curricula and delivery systems used in these experiments can be found in Head Start programs in many places and could easily be adopted by others, and while the children were typical of Head Start's populations, these were not typical Head Start programs. They were experimental programs. They varied in ages, frequency and duration of sessions. However, some were actually Head Start sponsored, and current Head Start quality standards are such that similar curricula are likely to be part of typical Head Start programs.

2. These studies were not initially designed for later comparisons or pooling of the data. This is a secondary analysis, and there are very real limits on the amount of information that was common across studies when the projects collected their initial baseline information. However, they are, with a few exceptions, all the existing studies that can be used for the investigation of long-term effects of early intervention. It would take another 15 years (and at least five million dollars) to create a similar sample.

3. The research questions in this report were somewhat different for these secondary analyses than those originally posed by the individual investigators. For example, we sometimes found it necessary to define treatment and control groups somewhat differently from the ways that the original investigators structured their comparison groups and analyses. Such changes are documented wherever they were made.

DESCRIPTION OF THE CONSTITUENT STUDIES

Although the various intervention programs had been independently conceived and carried out, they share several important characteristics:

1. All programs were initiated and for the most part completed prior to 1969.

2. The original samples were reasonably large, with a minimum size of 92 and a median size of 209.

3. Certain conditions of research design had been met.

4. Subjects had been followed after leaving the programs.

5. The intervention programs were explicit and standard so that the content of the children's experience could be specified.

Demographic data collected by the investigators at the time of the children's entry into the intervention programs[1] provide a rough picture of the characteristics of the populations from which the samples were drawn. The samples were similar in that the overwhelming majority of the subjects were Black (92%) and poor. Forty percent did not have a father living at home and 51% had three or more siblings. The children's mothers had completed a mean of 10.5 years of schooling. The mean household SES was 64.0 on the Hollingshead Two-Factor Index of Social Position, placing the mean in the lowest social class.[2] In 1976-77 during the follow-up data collection, subjects ranged in age from 9 to 19 years old.

---

[1] With the exception of Beller, who collected this data retrospectively.

[2] See the Consortium (1977) for a detailed comparison of samples across projects.

More specific data on the characteristics of the samples and the individual projects can be found in Tables 1 and 2.[1] The following descriptions of the projects include details of the programs offered to the children and the design of the research, including sample selection and assignment procedures. We also include our categorization of the projects, made for the purposes of the analyses included in this report as, either closely approximating an experimental design or a quasi-experimental design.

The Philadelphia Project: Dr. E. Kuno Beller

This program provided an experimental preschool for children who were between 3 1/2 and 4 1/2 years of age at entry. Four classes were established in four different public schools in North Philadelphia. Each class was staffed by an experienced head teacher and an assistant; the assistants were liberal arts graduates with no prior teaching experience.

Beller's nursery program could be classified as traditional in orientation. Its goals were diverse and included increasing the child's self-esteem as well as perceptual, cognitive, and physical development. It was "child-centered in...that...learning was shaped around the child's needs and preferences" (Beller, 1974).

---

[1] Background characteristics, IQ scores, and preschool attendance are listed only for the eight projects included in the regression analyses of school outcomes, to be detailed later in this report.

Besides the nursery subjects there were two other groups of children: children who entered regular kindergarten with no prior preschool experience, and those who entered first grade with neither preschool nor kindergarten experience. Beller's own research and analyses are concerned with comparing the effects of three different ages at entry to schooling.[1]

A pool of applicants was generated by sending letters to the parents of all students in four schools located in North Philadelphia, which is a predominately Black and very poor neighborhood. The nursery children were selected randomly from this pool of applicants. The kindergarten group:

> ...consisted of 53 5 year-olds who entered the same kindergarten (as the nursery children, but) without prior nursery experience. These children were selected from a larger group to approximate age, sex distribution, and ethnic background (of the nursery children) (Beller, 1974, p. 16).

The first-grade-only group was similarly selected from the first grade classes of the children in the first two groups.

Beller's first-grade-only group has been excluded from most analyses in this report in order to compare children who were more similar on background characteristics.[1] For the section of this report labelled "Non-cognitive Outcomes," however, Beller's first-grade-only and kindergarten groups were combined to create a single control group. Because of the procedures used in creating the three groups, Beller's study has been classified as quasi-experimental for this report.

---

[1] The research question in this report, however, was somewhat different for these secondary analyses. The treatment group was composed of children who attended experimental early education programs. Individual investigators, including Beller, may have structured their comparison groups and analyses differently.

Of Beller's original 112 subjects in groups one and two, 74 (66.1%) were included in the follow-up samples. For analyses of non-cognitive outcomes, 96 (56.4%) of the 170 subjects in groups one, two, and three were administered Youth Interviews. The modal age of these subjects at follow-up was 18. For additional information, see Beller, E.K., Impact of Early Education on Disadvantaged Children, in S. Ryan (Ed.), A Report on Longitudinal Evaluations of Preschool Programs. Volume I: Longitudinal Evaluations. Washington, D.C.: Office of Child Development, DHEW Publication No(OHD)74-24, 1974, pp. 15-48.

---

[1] Beller did not collect demographic data at the time of entry, but gathered retrospective demographic information while interviewing the mother at the time of follow-up. The first-grade-only group was significantly lower on mother's level of education. Five other retrospective variables did not reveal significant differences among the three groups.

Institute for Developmental Studies:

Drs. Cynthia and Martin Deutsch

The Institute for Developmental Studies, which was established in 1958, examined the effects of a curriculum developed especially by the Institute on several sample waves of children from low-income areas of New York City. The program focused on four general areas: language development, concept formation, perceptual and overall cognitive development, and the child's self-concept. Teaching methods and materials developed by the program were designed to help children master basic academic skills and become independent, confident learners. Children began the program during the prekindergarten year and continued through third grade.

Staff members actively recruited children for the program, obtaining names from a variety of community sources such as schools, churches and neighbors. Approximately one-third of the children, chosen randomly from the total sample, were designated as controls. This group of children first encountered formal schooling in the regular kindergarten classrooms of the New York City public school system.[1]

Although this project closely approximated an experimental design, it has been classified for this report as quasi-experimental because of problems of attrition. It was possible to retrieve only

---

[1] In order to provide additional controls, two other groups were formed, one at the beginning of each of the next two successive years. These control groups and later cohorts of children attending preschool were not included in the analyses reported here.

18.1% of the original 504 children in the first four waves of treatment/control children (Groups 1 and 2).

Severe Attrition Problems in the Deutsch Data

Unfortunately, we have had to omit the Deutsch project from most of the tables in this report because of the extent and nature of the attrition in that project. First, the attrition was vastly higher for the Deutsch project than for any other: the project with the median attrition rate (Palmer's) recovered 2.82 subjects for every subject lost, while the Deutsch project recovered only .221 subjects for every subject lost -- less than 1/12 as many.

Additionally, the Deutsch project recovered predominately the best treatment-group children and the worst control-group children. For example, for the school record data, the following occurred. At the end of the intervention program, when the children were aged about 5, the treatment group had a mean Stanford-Binet IQ score of 99.2 while the control group had a mean IQ score of 91.5. So far as we know, this difference of 7.7 points between mean IQ scores represented genuine (even if temporary) effects of the preschool program. However, when the same age-5 IQ scores are examined only for the sample whose school records were later recovered by the Deutsch team, the recovered treatment-group children had a mean age 5 IQ of 103.7 while the recovered control-group children had a mean age 5 IQ of only 84.2. Thus, if we had used only the recovered sample to look at mean IQ scores immediately upon completion of the preschool program, we would have observed a difference of 19.5 IQ points instead of the difference of 7.7 points which presumably represents real effects of preschool. Thus, using the recovered sample to examine more persistent effects would presumably have yielded biases of the same magnitude. This tendency to recover the best treatment-group subjects and the worst control-group subjects was statistically significant for the school record at the .001 level, by the interaction test described in this report's section on attrition. When the same interaction test was applied to the data from other projects, no biases nearly this large or significant were uncovered.

Several explanations might be given for these unfortunate biases. The Deutsch children were older than most others, and thus harder to find. Problems in New York City are often unlike those anywhere else in the nation. School systems were reorganized, school buildings torn down, records lost. Whatever the causes, it seemed most prudent to present the Deutsch analyses separately. See Appendix D for these results.

For additional information, see Deutsch, M., Taleporos, E., & Victor, J. A Brief Synopsis of an Initial Enrichment Program in Early Childhood. In S. Ryan (Ed.), A Report on Longitudinal Evaluations of Preschool Programs. Volume I: Longitudinal Evaluations. Washington, D.C.: Office of Child Development, DHEW Publication No (OHD)74-24, 1974, pp. 49-60.

The Parent Education Program: Dr. Ira Gordon

Dr. Gordon's project provided home-visitor, parent-focused intervention to children from 3 months to 3 years old in the Gainesville, Florida area. This study was specifically focused on the enhancement of the intellectual and personality development of the child and the production of changes in the mother's self-esteem and in her convictions that she could affect what happened to herself and her child. Gordon utilized trained paraprofessional home visitors who worked with each mother once a week. The sequenced curriculum emphasized Piagetian concepts appropriate to the child's stage of development. One treatment group received weekly visits for two years, starting when the child was 3 months old; a second, visits from 3 months to 1 year of age; and a third, visits from 1 year to 2 years of age. For the third year of the study, when children were 2 years old, a new treatment, termed the Home Learning Center,[1] was instituted. Children in groups of about five met in one family's home with their mothers and a paraprofessional teacher.

Gordon's assignment procedures were complex. Three waves of children were involved, and the assignment procedures varied among waves. All three waves were randomly assigned to treatment or control groups, but the assignment in one wave was not on an individual basis. That is, entire towns, (or, in the case of Gainesville, sections of town) were randomly designated as treatment or control areas. At the start of the second and third years, children were randomly re-assigned to treatment or no-treatment status. For example, for the Home Learning Center, a new group of 2 year olds was recruited and earlier participants were randomly assigned either to attend the small group experience or not. Thus, the experimental children received treatment for 1, 2 or 3 years and the control group received no treatment for the 3 year period. Because assignment was random, Gordon's project has been classified as experimental.

---

[1] Sometimes referred to as Backyard Centers.

Of the original 309 subjects, 107 (34.6%) were included in the follow-up. Gordon's subjects were 10 and 11 years old at the time of follow-up.

For further information about this project, see I.J. Gordon, The Florida parent education early intervention projects: A longitudinal look. Gainesville, Florida: Institute for Development of Human Resources, University of Florida, 1973. (ERIC Document Reproduction Service No. ED100 492.) (See also, Gordon, Guinagh, & Jester, 1977.)

The Early Training Project: Dr. Susan Gray

The Early Training Project, which was conducted in Murfreesboro, Tennessee, included both center-based and home-based components. It was explicitly concerned with fostering the children's intellectual growth and with their "attitudes related to school success" (Gray, 1974) and their general competence. In the summer programs, each group of about 20 children was served by one teacher and four assistant teachers; they typically worked in groups of five children with one adult. Traditional nursery school materials were employed, but their use was non-traditional in that activities were sequenced to become increasingly complex and were carefully focused on the goals of the program, e.g., increasing language use. The home visitor served as an active liaison between home and school during the summer periods. During the 9 month hiatus between summer programs, home visitors worked with each family once a week for a period of 1 hour in order to prevent erosion of gains made over the summer.

Sixty-five subjects were randomly assigned to one of two treatment groups or to one control group. The first treatment group participated in three summer (center-based) programs starting at age 4, with home visits spanning the 9 month intervals between summer programs. The program for the second treatment group was identical to that of the first, except that they entered the program in the second summer at age 5. In the analyses reported here, these two groups are merged into a single treatment group.

This report utilizes follow-up data on 55 of Gray's 65 subjects (83.6% of the original sample). This excludes data on a distal control group.[1] The modal age of the subjects at follow-up was 19.

For additional information, see Gray, S.W. Children From Three to Ten: The Early Training Project. In S. Ryan (Ed.),

---

[1] Gray recruited a second, self-selected comparison group in a nearby community, referred to as the "distal controls". The purpose of this group was to serve as a check against "horizontal diffusion" of treatment effects to non-treatment children in the same small community. The analyses reported here exclude this second comparison group in order to avoid weakening her strong experimental design. Additionally, distal controls were responsible for some differential attrition in this project.

A Report on Longitudinal Evaluations of Preschool Programs. Volume I: Longitudinal Evaluations. Washington, D.C.: Office of Child Development, DHEW Publication No. (OHD)74-24, 1974, pp. 61-68. Also Gray and Klaus, 1970.

Curriculum Comparison Study: Dr. Merle Karnes

In this study, two waves of children attended programs offering different curriculum models: Bereiter-Engelmann; Traditional; Community-Integrated; Montessori; and GOAL, Dr. Karnes' concept development curriculum. Each group attended one of the preschool models for about 2 hours a day for 7 to 8 months. The subjects were 4 year old children from families in the Champaign-Urbana, Illinois area who were classified as socio-economically deprived using Head Start guidelines.

The first wave was assigned to either a GOAL, Bereiter-Engelmann, or Traditional classroom; the second wave was assigned to either Bereiter-Engelmann, Montessori, or Community Integrated. There were no untreated controls. Classroom groups were stratified by IQ to insure a balanced range of intelligence scores in each class unit and to provide an opportunity to evaluate the effectiveness of the various programs on children from different ability groups. Children from higher IQ levels had to be actively recruited for the program. None of the children had IQ scores less than 70; the mean IQ score across all children was approximately 95. Class units were examined to assure comparability of sex and race. When necessary, substitutions were made between classes to maintain approximate ratios of 67% Black and 33% white plus 50% male, 50% female. Finally, each class unit was randomly assigned to a particular curriculum. Since Karnes did not select a comparison group, her data were not used in analyses of treatment/control differences.

Follow-up data are available on 88 of 102 original subjects (86.3%). At follow-up the subjects ranged from 11 to 16 years of age.

For additional information, see Karnes, M.B., Zehrbach, R.R., & Teska, J.A. The Karnes' Preschool Program: Rationale, Curricula Offerings and Follow-up Data. In S. Ryan (Ed.), A Report on Longitudinal Evaluations of Preschool Programs. Volume I: Longitudinal Evaluations. Washington, D.C.: Office of Child Development, DHEW Publication No.(OHD)74-24, 1974, pp. 95-104. See also, Karnes, Zehrbach and Teska, 1977.

Verbal Interaction Project: Dr. Phyllis Levenstein

The Verbal Interaction Project developed the Mother-Child Home Program, a home-based program for children aged 2 and 3 years and their mothers in an urban area on Long Island, New York. Commercially available books and toys were taken as gifts to the homes on a weekly basis by "Toy Demonstrators," who demonstrated techniques which were designed to encourage verbal interaction between the mothers and their infants. The toys and books were increasingly complex and were chosen to provide a structured cognitive curriculum. The program of 92 home sessions over two local school years was explicitly addressed to the mother and child as a socially interactive dyad. The overall aim of the program was to support the mother (and through her, the family) in fostering the intellectual and socioemotional development of her child. Other family members were encouraged to join the home sessions whenever possible.

Levenstein's selection and assignment procedures were quasi-experimental rather than experimental.[1] A number of treatment and control groups were created in three separate cohorts. Treatment groups varied in the amount of time they received home visits: 1 year, 18 months, or 2 years. In addition, an "after-only" control group was recruited in the first grade.[2] Follow-up data are available on 188 of 250 original subjects (85.2%); at follow-up, the subjects ranged from 9 to 13 years old.

For additional information, see Levenstein, P., The mother-child home program. In M.C. Day & R.C. Parker (Eds.), The preschool in action: Exploring early childhood programs (2nd edition). Boston: Allyn & Bacon, 1977. (See also, Madden, Levenstein, & Levenstein, 1976.)

Experimental Variation of Head Start Curricula: Dr. Louise Miller

Miller's study contrasted four types of Head Start curricula: traditional, Bereiter-Engelmann, Montessori, and DARCEE.[3] The

---

[1] Interestingly, 85% of mothers accepted when they were offered either treatment or control status.

[2] This control group was used only in the section entitled "Attitudes and Values."

[3] DARCEE stands for Demonstration and Research Center for Early Education. It was developed at George Peabody College in Nashville, Tennessee.

DARCEE subsample was further subdivided, with half receiving home visits and half not receiving home visits. Four target areas in Louisville were designated as appropriate sites for the program. Initially, each area was to receive all four programs, but a shortage of trained Montessori teachers resulted in only two instead of four Montessori classes. All other program types were carried out in each area.

A total of 214 4-year-olds were assigned to the experimental classes. Random assignment to classes was not used, as it would have entailed transporting children out of their neighborhoods. Rather, children were randomly assigned to programs within schools.

The selection of a comparison group proved problematic. An initial control group of 34 children from the same neighborhood -- including 21 children from the Head Start waiting list -- were selected at the outset of the program. However, this control group had a significantly higher percentage of whites and a significantly greater percentage of father-present homes than the treatment group at the outset; in addition, the control subjects had substantially higher mean incomes. This introduces an appreciable conservative bias into the calculation of treatment effects.[1]

---

[1] After the preschool year, some subjects from each group enrolled in Follow-Through kindergarten, while the remainder went into regular kindergarten. In addition, some of the Bereiter-Engelmann subjects went into a special Bereiter-Engelmann kindergarten. However, with one exception, none of the control children went into either Follow-Through or Bereiter-Engelmann kindergarten. Additional low-income children were recruited who had not been in preschool but who were in either the Follow-Through or the Bereiter-Engelmann kindergarten. These control children were excluded from most analyses with the exception of those in the "Non-cognitive Outcomes" section of the report.

Because children were not randomly assigned to control groups and control groups later proved to be significantly different from treatment groups on some background variables, Miller's project has been classified as quasi-experimental.

Follow-up data are available for 127 of Miller's original 248 subjects (51.2%). The modal age of the subjects at follow-up was 13.

For additional information, see Miller, L. & Dyer, J.L. Four Preschool Programs: Their Dimensions and Effects. Monographs of the Society for Research in Child Development, 1975, 40 (5-6), Serial No. 162.

Harlem Training Project: Dr. Francis Palmer

This program tested two models: a structured, concept training program[1] and a less structured "discovery" program. The Concept Training curriculum was designed to teach simple concepts believed to be prerequisite to subsequent learning. These concepts were introduced in order of increasing difficulty. The "discovery" program had no formal curriculum but children were otherwise treated identically to the concept training group (Palmer & Anderson, 1978).[2] Three durations of programs were tried: 123 children entered at age 2, 121 at age 3, with a subset of 20 children trained at both age 2 and 3. In all groups, children were brought to a center over a period of 8 months for twice-weekly sessions which lasted 45 minutes. Tutors worked one-to-one with each child.

---

[1] This curriculum, termd One-to-One, has subsequently been revised for home as well as center use.

[2] The children played with the same toys that were used to teach concepts, but tutors neither initiated conversations nor "taught" the child.

Treatment and control children were recruited from the birth records of the same two hospitals. The research design called for Black, male children from lower class and middle class families with certain birth criteria.[1] Children born in the months of August to October, 1964, were randomly assigned to a treatment group. Children born in November and December, 1964, were recruited specifically as controls. However, this selection procedure may not have introduced serious bias since the project staff emphasized the benefits of a total of 4 1/2 weeks of testing in recruiting the controls. One could thus see the control parents as volunteering for a less extensive program. Consequently, Palmer's study has been classified as experimental for the purposes of this report.

Follow-up data are avaiable for 228 (73.8%) of Palmer's initial 309 subjects. The modal age of the subjects at follow-up was 13.

For additional information, see Palmer, F.H. and Siegel, R.J., 1977. See also, Palmer and Semlear, 1976 and Palmer, Semlear and Fisher, 1978.

The Perry Preschool Project: Dr. David Weikart

The Perry program was a preschool program in Ypsilanti, Michigan. The children involved had all tested in the 50 - 85 range on an IQ test; their families had been rated as disadvantaged on an index which included educational and occupational levels and household density. Children entered at the age of 3 and attended half-day sessions 5 days a week for 2 school-years (October through

---

[1] Over 5 pounds at birth with English-speaking mother who had no history of drug addiction or venereal disease.

May). Teachers also made 90 minute weekly home visits. The curriculum was mainly Piagetian, focusing on cognitive objectives. Heavy emphasis was placed on experiences that would stimulate the child to construct concepts and develop logical modes of thought (Weikart, Deloria & Lawser, 1974).

Assignment to treatment and control groups was essentially random. The only not worthy exception to this generalization is that in cases where a child assigned to the treatment group could not attend due to lack of transportation or maternal employment (preventing scheduling of home visits), the child was exchanged with a matched child assigned to the control group. Unfortunately, this produced a group difference on maternal employment:[1] 8.6% of the mothers in the experimental group vs. 30.8% of the control mothers, were employed (p = .002, two-tailed). Nonetheless, this study has been classified as experimental as this one exception was deemed a relatively minor departure from a pure experimental design.

Four cohorts were included in the study. At the time of follow-up, the subjects ranged from 15 to 19 years of age. Follow-up data are available on all of the 123 original subjects.

For additional information, see Weikart, D.P., Deloria, D.J., & Lawser S. Results of a Preschool Intervention Project. In S. Ryan (Ed.), A Report on Longitudinal Evaluations of Preschool Programs. Volume I: Longitudinal Evaluations. Washington, D.C.: Office of

---

[1] Subsequent analyses of maternal employment when children were aged 15 and analyses of whether mothers had ever been employed found no significant differences between treatment and control children.

Child Development, DHEW Publication No. (OHD)74-24, 1974, pp. 125-133. See also, Weber, C.U., Foster, P.W., & Weikart, D.P. (1978); Weikart, D.P., Bond, J.T., & McNeil, J.T. (1978); and Weikart, D., Deloria, D., & Lawser, S. (1970).

Micro-Social Learning System: Dr. Myron Woolman[1]

Dr. Woolman studied the effects of a preschool program utilizing an arrangement of modular units in which children worked through a pre-planned series of activities. They received periodic reinforcement as they completed each objective in a sequence. The program design also included a life-simulator space in which the children applied their newly learned skills in free play. This aspect of the program utilized materials and equipment designed to provide unstructured free response favoring interactive play.

The program group, selected by the Vineland school district, consisted primarily of the children of migrant families, the majority of whom were on welfare. About 10% of the group were children of higher socioeconomic status whose parents had requested that they be allowed to enter the program. Since the program group consisted of the highest risk children in the school district, there did not exist a sufficiently large non-treated group which could be used as a control group. Therefore, it was decided to compare the program children to the general school population. A random sample of the previous year's first grade population, including both middle and lower class children, was selected.

---

[1] Note: Since Dr. Woolman's data were complete as of July, 1977 this group was not included in the October, 1977 - August, 1978 analyses. For a report of the effectiveness of this program, see the Consortium, 1977. Briefly, the program significantly reduced the number of children retained in grade.

Because the program group had a much higher percentage of Spanish-surnamed children than the general school population, an additional 36 Spanish children were randomly selected from the prior year's first grade population. This group allowed the additional comparison of Spanish-surnamed program children to Spanish-surnamed children in the general school population and, in effect, constituted a control for ethnicity and social class.

For additional information, see Woolman, M., Learning for Cognition: The Micro-Social Learning System, Report to the New Jersey State Department of Education, 1971.

New Haven Follow Through Study: Dr. Edward Zigler

This study investigated the effects of Head Start and Follow-Through programs on two cohorts of children in New Haven, Connecticut. The original group has now been followed through the eighth grade. In the first wave the program groups consisted of children recruited for Follow-Through in several low-income areas. The control group consisted of economically disadvantaged children in one classroom from each of three schools located in similar low-income areas. For the second wave, the control group consisted of children drawn randomly from the same schools from which the original controls had been drawn. Since the program and control children were not drawn from a common pool of children whose parents had volunteered them for Follow-Through, the extent of the self-selection bias in unknown. It is likely to be minimized by the fact that the controls were drawn from different schools than the program children; that is, the controls did not consist of children whose parents had decided not to participate in the program. Therefore, this project has been classified as quasi-experimental.

At follow-up, 185 of the original subjects were retrieved. The modal subject was aged 13 years.

For further information, see Seitz, V., Apfel, N., & Efron, C. Long-term Effects of Early Intervention: The New Haven Project. In B. Brown (Ed.), Found: Long-term Gains from Early Intervention. Boulder, Colorado: Westview Press, 1978. (See also, Abelson, Zigler, & DeBlasi, 1974.)

Table 1

Characteristics of Early Education Programs and Ages of Subjects for Each Data Set

Principal Investigator	Early Education Program	Location	Population (1970) 000	Type of Delivery System	Subject Birth Year	Age at Entry to Program	Length of Program (years)	Years of Program	Age at 1977 Follow-up
Beller	The Philadelphia Project	Philadelphia, Pennsylvania	1,949	Center	1959	4 years	1 year	1963-64	18
Deutsch	Institute for Developmental Studies	New York, New York	7,895	Center	1958-1962	4 years	5 years	1963-71	15-19
Gordon	The Parent Education Program	Gainesville, Florida	64	Home	1966-1967	3 mos. to 2 yrs.	3 years	1966-70	10-11
Gray	The Early Training Project	Murfreesboro, Tennessee	26	Home/ Center	1958	3.8 or 4.8 yrs.	14 mos. or 26 mos.	1962-65	19
Karnes	Curriculum Comparison Study	Champaign-Urbana, Illinois	89	Center	1961-1963	4 years	1 year	1965-66 (2 waves)	14-16
Levenstein	The Mother-Child Home Program	Glen Cove, Manhasset and Freeport, Long Island, New York	26 8 40	Home	1964-1968	2 yrs. & 3 yrs.	1 - 1 1/2 years	1967-72	9-13

Table 1 (Cont.)

Principal Investigator	Early Education Program	Location	Population (1970) 000	Type of Delivery System	Subject Birth Year	Age at Entry to Program	Length of Program (years)	Years of Program	Age at 1977 Follow-up
Miller	Experimental Variation of Head Start Curricula	Louisville, Kentucky	361	Center & Center/ Home	1964	4 years	1 year	1968-69	13
Palmer	Harlem Training Project	New York, New York	7,895	Center	1964	2 or 3 years	1 or 2 years	1966-68	13
Weikart	Perry Pre-school Project	Ypsilanti, Michigan	30	Center/ Home	1958-1962	3 yrs. (1st wave) 4 yrs.	2 yrs. (1st wave) 1 year	1962-67 (5 waves)	15-19
Woolman	Micro-Social Learning System	Vineland, New Jersey	47	Center	1963-1968	4-5 yrs.	1-4 yrs.	1969-73	9-14
Zigler	New Haven Follow-Through Study	New Haven, Connecticut	138	Center	1962	5 years	4 years	1967-71	15

Table 2

Background Characteristics, IQ Scores, and Preschool Attendance for Each Data Set

Data Set (n)[a]	Mean Mother's Educational Level	Mean No. of Siblings	Mean Pretest IQ Score	Mean IQ Score at 6 Yrs.	Percent Father Present	Percent Black	Percent Male	Percent Preschool Participants (vs. control)
Beller (56)	10.94 (1.7)[b]	2.96 (2.4)[b]	92.89 (55)[c]	97.25 (53)[c]	75.0	92.9	50.0	58.9
Gordon (64)	9.98 (1.9)	2.59 (1.9)	----	92.50 (62)	---	92.2	43.8	89.1
Gray (52)	8.67 (2.7)	4.17 (2.3)	89.25 (48)	90.94 (50)	67.3	100.0	50.0	65.4
Karnes (61)	10.16 (1.9)	3.46 (2.7)	95.84 (61)	104.75 (56)	62.3	62.3	50.0	100.0
Levenstein (125)	10.69 (1.9)	2.50 (1.6)	84.52 (121)	97.67 (118)	70.4	94.4	56.8	81.6
Miller (120)	10.68 (2.0)	3.24 (2.1)	----	94.42 (120)	45.0	91.97	46.7	85.0
Palmer (219)	11.13 (1.8)	2.41 (2.1)	92.12 (132)	95.54 (195)	72.6	100.0	100.0	78.5
Weikart (123)	9.42 (2.2)	3.89 (2.6)	79.02 (123)	88.63 (120)	52.9	100.0	58.5	47.2
Mean (820)	10.21	3.15	89.82[d]	94.86[d]	63.6	91.7	50.8	72.2

Note. IQ scores are Stanford Binet (except PPVT for Levenstein). Palmer IQ scores are at age 5 instead of 6. Data are not available for Gordon and Miller pretest IQ score, and Gordon father presence.

[a] Figures in parentheses indicate number of children in all calculations except IQ scores.

[b] Figures in parentheses below mean mother's educational level and mean number of siblings are standard deviations.

[c] Figures in parentheses indicate number of children in IQ score calculations.

[d] Mean IQ scores exclude Levenstein PPVT scores.

METHODS

The research reported here and in previous publications represents the results of secondary analysis of samples of children who have been followed for a number of years. This section of the report will discuss some of the issues and problems which arise in the course of such analysis and will outline the nature of the solutions used by the central staff.

General Problems in Longitudinal Assessments of Program Effects

Longitudinal data, such as that collected by Consortium members, provide a valid and direct way of assessing the cognitive, social, emotional and familial outcomes of programs for young children. The Consortium studies are further strengthened by the use of control or comparison groups. An alternate strategy -- using children as their own controls -- inextricably confounds program effects with maturational changes which would occur regardless of experience in programs. However, these strengths are accompanied by attendant weaknesses: issues of equivalence of treatment and control groups and problems of attrition and measurement.

The dangers of non-random assignment to treatment and control groups and the general inadequacy of standard statistical techniques that "correct for" sampling biases have been well-documented in evaluation literature (e.g., Campbell & Erlebacher, 1970). If random assignment is not used, there is no way to know whether the treatment and control groups were initially comparable in all respects, and, hence, no way to accurately assess whether the treatment had any effect. Doubts can be raised even when numerous indicators of initial status (e.g., IQ, SES) were collected at the outset since it can always be hypothesized that the "true difference" between the treatment and control groups was not reflected in these measures. Self selection of either subjects or controls is also reason for concern. When subjects are self-selected, their later outcomes may reflect characteristics which led them to volunteer (to be either subjects or controls) rather than characteristics of the treatment.

In general, randomization is difficult to attain in evaluation research for both practical and ethical reasons. However, some projects in the Consortium more nearly approximate experimental designs than others. For purposes of data analysis, we have designated certain projects as experimental and others as quasi-experimental based on their sample selection and assignment procedures.[1] This procedure is intended as a safeguard to help insure that the findings are not artifacts of limitations in individual studies. Our categorization, however, is open to question and other reviewers might reach different decisions.

---

[1] Those designations were as follows. Experimental: Gordon, Gray, Palmer, Weikart. Quasi-experimental: Beller, Deutsch, Levenstein, Miller, Woolman, Zigler.

The Problem of Attrition

When one of our co-directors (Darlington) joined this project, he was, like many people, extremely suspicious of the problem of attrition in the Consortium data. We have subsequently analyzed the problem of attrition in an unusually thorough manner, performing hundreds of significance tests, any one of which might indicate some problem involving attrition. These tests concern the following four questions: (a) What is the overall rate of subject loss? (b) Are the subjects who have been lost different in important ways from those who have been retrieved? (c) Is the rate of subject loss different for treatment and control groups? (d) Do the characteristics of the lost subjects differ as a function of group membership (for example, did the study lose the brightest controls and the low-IQ treatment subjects)? These analyses did reveal some rather severe problems concerning attrition in the Deutsch data. As a result, the Deutsch data have been omitted from all our major analyses. These problems are described more fully in the section on project descriptions. In data from the remaining projects, our analysis supported the view that the final samples are representative of the original sample and the final sample treatment and control groups appear to be equivalent. The complete attrition analysis is presented in Appendix A.

Problems of Measurement

Longitudinal research also poses problems of measurement, one of which is particularly relevant to the present report: the potential non-equivalence of identical measures administered at different ages. The data on which the present report is based are derived from essentially identical tests and interviews administered to subjects ranging in age from 9 to 19. Since treatment subjects were compared to control subjects of similar ages, there is no reason to expect that age-related non-equivalence of the instruments would produce artifactual treatment-control differences. Indeed, such non-equivalence would be expected to attenuate true treatment effects by contributing additional random variation within the treatment and control samples. However, it could be a serious problem in several other types of analysis, such as age main effects, age (at measurement) by treatment interactions, and contrasts between programs (as the programs differ in age at follow-up).

An additional problem of longitudinal research is what Campbell (1971) and others have called "temporal erosion". While a developmentalist might quarrel with the simplicity of Campbell's statement that "all relationships weaken with time", it would certainly be expected that many, if not most, treatment-control differences would erode with time. In the present report, the lag between the end of the subjects' intervention experience and the time of follow-up data collection is as much as 14 years -- time enough for even very gradual erosion to take a heavy toll. On the other hand, it is possible that intervention programs not only taught concrete skills, but altered the context in which the children operated. For example, the mother-child interactive system may have been affected by intervention and this "new" interactive system could operate in such a way that erosion would be lessened. More concretely, suppose a given program had an effect on cognitive abilities and on school performance which, while strong, was relatively shortlived (2 years, let us say). The initial improvement in the child's school performance might affect the mothers' and the teachers' expectancies and their behavior towards the child. This could, in turn, affect the child's own attitudes and behavior. Thus, the simple original "treatment" could reverberate through time.

Campbell (1971) recommends the use of repeated measurements in order to assess temporal erosion and to gauge treatment effects realistically. Lacking that opportunity in the present case, we can only be alerted to the issue and keep in mind that the data presented here test only the null hypothesis of no persistent effects. A failure to disconfirm the null hypothesis could be entirely due to the non-persistence of effects rather than an initial absence of effects.

Issues Inherent in Multi-sample Secondary Analysis

The data presented here include a number of different programs (and samples) and hence, in most cases, can be seen as a set of replications of each hypothesis test. A number of issues arise in choosing how to deal with data of this sort.

The first choice involves the unit of analysis: should analyses consider each subject in each program as an individual case? There are some instances (e.g., the analysis of differential effectiveness of specific preschool programs; see Koretz, Vopava and Darlington, 1978) where the only valid option is to use programs as the unit. In others, the child is naturally the unit of analysis; that is, when the independent variables under study are properties of the child such as family background, IQ scores, etc.

When using subjects as a unit, the simplest approach would be to pool all subjects into a single large sample prior to analysis. This approach, however, is unacceptable because differences between samples can yield artifactual results.[1] Instead, unless otherwise noted, we have pooled subjects within projects only.

The next issue is how to pool results across samples. Unless otherwise noted, we have used a method whereby $\underline{z}$ scores are summed. First, the exact $\underline{p}$ value of results from each sample is converted to a $\underline{z}$ and given a sign according to the direction of the effect. These $\underline{z}$ scores are then summed by the formula:

$$z = \frac{\Sigma z_i}{\sqrt{k}},$$

---

[1] As a simple example, consider an analysis which includes two totally ineffective programs. Let program A have a treatment group of size 100 and a control group of size 20. Let both groups have mean pre- and post-test IQ scores of 90. Program B has 20 treatments and 100 controls, with pre- and post-test means of 80. Analyses of each group separately would correctly show no treatment effect. If the two samples were pooled, however, the resulting sample would show a post-test mean difference of 6.7 points.

where z_i is the *z* score from sample i and k is the number of samples (Mosteller & Bush, 1954).[1] This method is more powerful than an analogous chi-square technique (c.f. Mosteller & Bush, op. cit.) for cases in which results tend to be consistently small but are in the same direction. Additionally, it is directional, and contrary results in different samples will be cancelled out. Thus, this formula can be seen as testing the presence of an overall 'average' effect.[2]

While the method above tests the magnitude of an average, overall effect, one might also want to test the consistency of a result across programs. To do so, programs are treated as the unit of analysis.

Analyses based on programs as the unit of analysis involve far smaller sample sizes than those based on subjects, and many people would assume that the decrease in sample size brings with it a commensurate decrease in power. However, as Darlington (1978a) points out, this is not necessarily the case, due primarily to the greater stability of group means as compared to individual scores.

---

[1] See also, Darlington, 1975.

[2] One null hypothesis with regard to preschool effects on later outcomes might be that no program had any effect. In this case, finding even one program with significant effects would be enough to reject the null hypothesis. However, the null hypothesis we chose to investigate is: averaged across many programs, preschool does not affect later outcomes. In this case, if three programs were to have a positive effect and three a negative effect, they would cancel each other out and, as a result, the null hypothesis could not be rejected.

A study on a few dozen observations, each of which is very stable because most of the variation caused by single individuals has been averaged out, can be nearly as powerful as one with 2,000 observations. This effect is not just a theoretical one; we have observed a number of times in our data that significance levels computed with the smaller N are of comparable magnitude to those computed with the much larger N. Yet the smaller N is far more valid a procedure, because it handles the problem of nonindependent observations (e.g., the effects of an unusually skilled teacher across one classroom).

An additional concern in pooling results, closely related to the issue of the consistency of results across samples, is the robustness of the test. As we shall use the term here, "robustness" refers to the ability of an overall significant result to withstand the removal of one or more samples. An overall significant result might be caused by exceptional results from one project or program. In field research like this, it is impossible to eliminate all the experimental inelegancies that can be eliminated in true laboratory research, and we have to assume that in a dozen projects, there must have been known or unknown design problems in at least one or two -- and we don't know which one or two. To take this possibility into account, we rank-order the projects or programs in terms of the degree to which the results from the project or program confirm our experimental hypothesis. We then delete from the analysis, one by one, the projects or programs whose data most strongly confirm our experimental hypothesis. If a significant result evaporates after just one such deletion, we do not consider the finding a strong one.

Finally, the reader will note that analyses throughout this report are performed separately for varying numbers of subgroups and frequently utilize many different dependent measures. Because of the large number of tests, it would be possible to obtain "significant" results by chance alone. To correct for this possibility, we have, as a general policy, used the Bonferroni/Ryan statistical technique to correct for multiple comparisons. This technique involves multiplying the significance level of each result by the number of tests performed (c.f. Darlington, 1978b). This is a conservative procedure in that all significant results are not accepted at face value; but it maximizes our confidence in the results.

COGNITIVE EFFECTS OF PRESCHOOL

Intelligence Test Scores

For reasons of availability and comparability, many intervention programs routinely administered IQ tests to participating children. The Consortium projects had information on children's IQ measured before and soon after intervention. In addition, the 1976-77 follow-up included WISC-R IQ tests. As the Consortium (1977) reported, early education proved to have a significant effect on children's IQ scores (Stanford-Binet) that lasted at least 3 years beyond the preschool experience. After analysis of the WISC-R IQ data, however, the investigators concluded that the effect of preschool intervention programs on children's IQ was not permanent, since projects with children older than 10 years showed no statistically significant treatment/control differences.

Those initial findings have now been extended by: (a) evaluating the effects of preschool experience, independent of family background variables and initial IQ, on post treatment IQ scores at age 6;[1] (b) performing analyses that include additional WISC-R data received since July, 1977; (c) investigating the effects of preschool on each of the WISC subtests; and (d) investigating the effects of preschool on the variability of WISC-R at later ages.

The first question is very important since it is possible that the earlier report of preschool effectiveness at raising IQ scores for 3 years was due primarily to factors other than the preschool experience. Questions (b), (c), and (d) above are important in order to understand in more detail the long-term impact of preschool on IQ.

---

[1] See the list of Consortium members' reports and publications for other treatments of this question.

Methods

Although the various projects comprising the Consortium were initiated at different times and were designed for children of a range of ages, nearly all had administered individual Stanford-Binet IQ tests to their subjects when they were 6 years old.[1] In order to help insure that the IQ gains found earlier for children who attended preschool were not caused by initial (pretreatment) differences between the treatment and control group, multiple regression was used. This technique allows one to assess the independent effect of preschool attendance on IQ at age 6 [question (a) above]. The independent variables in the regression were: child's sex, initial IQ score, family structure (father absent or present), family size (number of siblings) and mother's level of education. The unit of analysis was the individual child within a project.[2] Analyses were performed for each project. The *F* value for the regression coefficient for preschool vs.

---

[1] The IQ data used differed for two projects. Palmer's IQ scores are for children at age 5. Levenstein's project had more data for the Peabody Picture Vocabulary Test than for the Stanford-Binet test. Accordingly, the pretest and posttest IQ scores for Levenstein are PPVT scores.

[2] The *N* for each project may differ somewhat from earlier reports or from figures cited in this report. For these analyses, only children with complete information for every independent variable were used.

control was converted to a z score and z scores were then pooled and tested for significance according to procedures described earlier.

For the WISC-R analyses [questions (b), (c) and (d) above], each project was considered separately. Data were not merged across projects and significance levels were not combined across projects. This procedure was followed because the WISC-R tests were administered to subjects of different ages across projects and different lengths of time had elapsed since the preschool experience. In effect, this meant that the projects were not asking the same research question. Comparisons across projects for these questions, then, are only descriptive.

Results

Five data sets[1] had sufficiently complete data to test the simultaneous predictive powers of pretest IQ, family background measures, and preschool attendance on children's Stanford-Binet scores at age 6. As Table 3 shows, preschool attendance contributed to an increase in IQ scores independent of the effects of initial IQ level, child's sex and three measures of family background. The finding was very strong, both for the experimental data sets ($p < .0001$), and all five data sets ($p < .0001$). It was also robust; when the two strongest findings (Gray and Levenstein) were deleted, the result remained highly significant ($p = .005$).

---

[1] Because pretest IQ scores were not available for the Miller and Gordon data sets, they were not included. Miller administered initial IQ tests after treatment children had been enrolled in preschool for 6 weeks. Gordon's children were below age 2 at the time of first testing; infant test scores are not comparable to Stanford-Binet scores (Lewis & McGurk, 1972). Palmer's children who began the program at age 2 do not have a pretest IQ score.

66

Another regression analysis was performed, including the Miller and Gordon data sets, to assess the effects of sex, family structure, family size, maternal education and preschool attendance on IQ scores at age 6. Again, preschool made an independent contribution to IQ at age 6, as shown in Table 4. This result was significant both for the experimental data sets ($p < .0001$, $N = 433$) and for all seven data sets ($p < .0001$, $N = 730$).[1] This finding was also robust; when the two data sets with the strongest p values were deleted (Gray and Levenstein), the result remained significant ($p = .0086$).

In summary, these two analyses indicate that the increase in IQ scores at age 6 shown by children who had participated in preschool programs was attributable to the preschool experience, independent of the effects of sex, initial IQ, and various measures of family background.

Next we turn to a consideration of children's IQ scores as measured by the WISC-R when the children were aged 9 to 19 years. Mean IQ scores of treatment vs. control groups were compared using t-tests. As Table 5 shows, the additional information based on data received since July, 1977 does not alter the earlier published report that effects persist for several years but are not permanent.

---

[1] Three data sets did not find significant results: Miller, Gordon, and Beller. Miller's control children had somewhat higher IQ scores compared to her Head Start children, but the difference was not significant. The Beller and Gordon treatment children had higher IQ scores at age 6 than their respective controls, but the difference was not significant in the subset of children used in this study.

As reported earlier (Consortium, 1977), Levenstein's treatment group surpassed the control group on the full-scale ($p = .002$); verbal ($p = .0001$) and performance ($p = .063$) scores. Palmer's treatment group scored higher than controls on the performance score ($p = .041$). The only project added to the analysis since that report (Gordon) showed no significant differences on any of the IQ measures.[1]

T-tests were also performed on each of the WISC-R subtests to compare scores of program and control children. Table 6 shows that generally no subtest differences appeared in projects where there were no IQ differences. For three projects, Gordon, Gray, and Weikart, there were no significant differences at all. The Miller project revealed no significant differences when corrected for multiple comparisons (i.e., multiplying the p value by the number of subtests). Only the Levenstein and Palmer projects revealed statistically significant differences on subtests (after correcting for multiple comparisons). These differences were reflected in significant overall IQ differences (see Table 5).

Variability of WISC-R IQ scores was investigated because it seemed plausible that early intervention might help some students but hurt others -- for example, by making their ordinary public school years seem drab by comparison. This hypothesis was tested by an F test. As Table 7 shows, only the Verbal IQ of Gray's subjects and the Full Scale IQ of Palmer's subjects showed significant differences in standard deviation. In both cases the controls had significantly larger standard deviations than the program groups.

---

[1] Gordon and Guinagh (1978) reported that children who had received all 3 years of the program, the first 2 years, or the second 2 years of the program had significantly higher WISC IQ scores than control children at age 10. Our analysis combined into one group children with 1, 2, or 3 years of the program anc compared them to the controls. The overall results in our analyses for Gordon's project found no significant differences.

Table 8 shows results of the same analysis for the WISC subtests. Of the 65 comparisons of subtest standard deviations, only six showed significant differences, and this number was reduced to three after correcting for multiple comparisons: Gray (comprehension), Palmer (information), and Weikart (coding). Two of the three (Gray and Palmer) showed greater standard deviations for controls, one (Weikart) showed the opposite. Although the individual significant differences could be investigated, our overall conclusion is that there is no evidence for a general effect of preschool on the variance of any WISC-R full IQ or subtest score.

In summary, only the Palmer and Levenstein projects, found treatment/control differences on IQ at follow-up. Gordon's own analyses found treatment/control differences among subjects with more than 1 year of the program. All these children were below 13 years of age. In projects with children aged 13 years and above, there were no treatment/control differences on either the full-scale WISC scores or on the WISC subtests. After correcting for multiple tests, there were also no significant differences between standard deviations of the IQ scores of treatment and control groups.

On the basis of similar findings (that preschool did not permanently raise children's IQ scores compared to control populations) the usefulness of intervention programs was questioned and funds were frozen so that federally-sponsored programs could not expand. Therefore, it is important to consider what the current findings mean.

As the Consortium (1977) reported, preschool attendance was associated with higher IQ scores for at least 3 years after the programs had ended. While an increase in IQ score is usually interpreted as a gain in cognitive understanding and ability, it may also reflect some changed behavior patterns such as more spontaneous verbalization, less distractability, greater task-orientation, more ability to cooperate with adult demands and to adapt to structured situations (cf. Hertzig, et al., 1968; Moriarity, 1961). The reader will notice that all the non-cognitive gains which could help raise an IQ score are also relevant for learning and performing in school. Schweinhart and Weikart (1978) propose the following explanation:

> ...preschool improves intellectual functioning and adaptive behavior. The improvement in intellectual functioning fades away, but the improvement in adaptive behavior remains and leads to improved academic achievement throughout the school years. We offer this explanation not as a final answer, but as a line of thought that may be worthy of pursuit (p. 27).

Because the intelligence test has been with us for 70 years, it is sometimes easy to forget that it is merely an operational measure of "intelligence". "Intelligence" is not some bounded entity with a clear causal relationship to performance but a conceptual representation of abilities considered necessary for adequate functioning. But functioning within which domain? One thing that the IQ test does very well is to predict school grades. Rather than use IQ scores as predictors of school performance, however, it would be more useful to examine grades themselves and other direct indicators of school performance.[1]

What does the IQ score mean for everyday coping ability? Mercer (1975) has shown that the same IQ score may have different meaning depending on the cultural background of the individual. Lower class Blacks and Chicanos with IQs below 70 were more able to perform everyday chores and thus live independently than were middle class whites with the same low scores. Thus, in these senses at least, IQ scores may be said to have limited usefulness. In Bronfenbrenner's (in press) terms, there are other "more ecologically valid methods for assessing development-in-context." One such method is to gather evidence of whether individuals adapt to and cope with the demands made upon them by specific institutions, e.g., the school. The section which follows contains such an analysis.

---

[1] As McClelland (1973) points out, 15 years later it is more important to know what level of schooling the individual attained rather than whether or not s/he made A's. And the correlation between IQ scores and level of vocational success is more than likely an artifact of their joint association with social class.

Table 3

The Effect of Early Education on IQ Score at Age 6 When Background Variables and Pretest IQ Are Controlled

Data Set (n)	Preschool Coeff.	F	Signif. Level (2-tailed)	Pooled z Score	Pooled p (2-tailed)
Approximately Experimental Design					
Gray (47)	12.971[a]	29.271[b]	<.0001		
Palmer (118)[d]	1.532	.582	.4482		
Weikart (120)	4.223	5.559	.0204		
Experimental Total (285)				4.4572[c]	<.0001
Quasi-Experimental					
Beller (57)	4.547	3.565	.0661		
Levenstein (114)	9.961	8.969	.0035		
All Data Sets (456)				5.5812	<.0001

Note. Equation: IQ6 = IQPRE + MED + FP + SIBS + SEX + PC.

Legend: IQ6 = IQ score at six years old; IQPRE = pretest IQ score; MED = mother's education; FP = father presence; SIBS = number of siblings; PC = preschool vs. control.

[a] Unstandardized regression coefficient for preschool vs. control.

[b] F test for significance of preschool coefficient, simultaneous (standard regression) method. Figures in table are as if preschool were added after all other variables.

[c] Pooled $z = \frac{\Sigma z_i}{\sqrt{k}}$ z_i = z score from sample i and k is number of data sets.

[d] The Palmer project administered the Stanford-Binet IQ test at age 5.

Table 4

The Effect of Early Education on IQ Score at Age 6 When Background Variables Alone Are Controlled

Data Set (n)	Pre-school Coeff.	F	Signif. Level (2-tailed)	Pooled z Score	Pooled p (2-tailed)
Approximately Experimental Design					
Gordon (62)	7.090[a]	1.809[b]	.1858		
Gray (50)	13.462	12.582	.0010		
Palmer (210)[e]	5.796	7.025	.0088		
Weikart (120)	4.833	6.058	.0156		
Experimental Total (433)				4.8283[d]	<.0001
Quasi-Experimental					
Beller (57)	2.194	.365	.5503		
Levenstein (118)	9.530	7.757	.0064		
Miller (122)	-2.419	.688	(.4095)[c]		
All Data Sets (730)				4.5945	<.0001

Note. Equation: IQ6 = MED + FP + SIBS + SEX + PC. For legend, see Table 3.

[a] Unstandardized regression coefficient for preschool vs. control.

[b] F test for significance of preschool coefficient, simultaneous (standard regression) method. Figures in table are as if preschool were added after all other variables.

[c] Figures in parentheses are in reverse direction -- preschool had a negative association.

[d] Pooled $z = \frac{\Sigma z_i}{\sqrt{k}}$ See Table 3.

[e] The Palmer project administered the Stanford-Binet IQ test at age 5.

Table 5

Mean WISC IQ Scores, Program vs. Control by Project

Project		Mean Age	Mean IQ Program	Mean IQ Control	*t* Value	Significance	*N* Program	*N* Control
Gordon	FIQ	10-4	83.11	79.05	1.26	.210	70	20
	VIQ	10-4	82.97	78.15	1.52	.132	70	20
	PIQ	10-4	85.59	83.35	0.68	.499	70	20
Gray	FIQ	16-9	78.74	76.44	0.59	.558	34	18
	VIQ	16-9	77.09	76.61	0.12	.909	34	18
	PIQ	16-9	83.91	79.44	0.94	.351	34	18
Levenstein	FIQ	9-9	101.86	93.56	3.21	.002 *	51	25
	VIQ	9-9	98.41	89.36	3.66	<.001*	51	25
	PIQ	9-9	105.45	99.48	1.89	.063	51	25
Miller	FIQ	12-8	84.96	87.69	1.13	.262	109	32
	VIQ	12-8	83.08	85.53	0.96	.337	109	32
	PIQ	12-8	89.42	92.41	1.12	.263	109	32
Palmer	FIQ	12-2	92.13	88.86	0.99	.327	104	28
	VIQ	12-2	93.36	91.25	0.74	.461	104	28
	PIQ	12-2	92.34	86.39	2.07	.041*	104	28
Weikart	FIQ	14-0	81.02	80.71	0.14	.885	54	56
	VIQ	14-0	78.33	77.64	0.36	.721	54	56
	PIQ	14-0	87.59	87.82	0.10	.924	54	56

---

* $p < .05$ before correction for multiple comparisons.

Table 6

Mean WISC Subtest Scores, Program vs. Control by Project

Project	Subtest$^{\pm}$												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Gordon													
Control	6.11	5.84	7.79	5.47	7.37	--	8.63	8.11	6.68	6.53	--	7.63	--
Program	6.71	6.78	8.53	5.93	7.75	--	9.21	8.03	6.69	7.50	--	7.56	--
_t_	.93	1.34	1.06	.66	.57	--	.79	.08	.01	1.46	--	.10	--
Significance	.353	.185	.290	.510	.572	--	.430	.933	.993	.148	--	.918	--
Gray													
Control	5.78	6.17	7.11	5.83	6.17	--	6.72	7.78	5.83	7.06	8.00	--	--
Program	5.91	6.26	6.94	6.24	6.15	--	7.38	7.18	6.97	8.18	8.24	--	--
_t_	0.19	0.13	0.22	0.56	0.02	--	0.74	0.61	1.39	1.12	0.24	--	--
Significance	.849	.894	.827	.579	.981	--	.465	.544	.169	.270	.814	--	--
Levenstein													
Control	6.72	9.60	8.08	7.68	8.16	9.60	9.04	10.44	8.80	9.72	--	11.64	--
Program	8.57	11.00	9.94	9.20	9.24	10.18	9.75	10.10	10.10	11.16	--	12.82	--
_t_	3.32	2.30	3.18	2.41	1.92	.96	1.27	.56	2.17	1.92	--	1.42	--
Significance	.001*	.024*	.002*	.018*	.059	.338	.207	.579	.034*	.058	--	.160	--

Table 6 (Cont.)

Project	Subtest 1	2	3	4	5	6	7	8	9	10	11	12	13
Miller													
Control	6.75	7.84	8.09	6.72	9.06	7.31	8.94	9.28	7.44	9.00	--	10.03	9.09
Program	6.51	7.31	7.86	6.35	7.86	7.58	8.75	8.76	7.90	8.17	--	8.96	8.95
_t_	.52	.86	.49	.64	2.04	.39	.36	.89	.59	1.27	--	1.60	.23
Significance	.601	.391	.624	.522	.043*	.694	.721	.377	.554	.205	--	.112	.816
Palmer													
Control	7.82	8.57	8.82	8.29	10.18	8.54	9.00	8.82	7.82	7.32	--	6.39	10.04
Program	8.40	8.54	9.18	8.55	10.08	9.09	9.83	9.75	8.31	8.53	--	8.51	10.17
_t_	.52	.86	.49	.64	2.04	.39	.36	.89	.59	1.27	--	1.60	.23
Significance	.455	.961	.550	.679	.870	.353	.122	.090	.423	.100	--	.002*	.841
Weikart													
Control	5.59	8.36	5.89	5.64	6.77	--	7.54	7.93	7.59	7.75	--	9.95	--
Program	5.48	8.15	6.31	5.93	6.67	--	7.61	7.56	7.70	8.00	--	10.04	--
_t_	.32	.55	1.08	.66	.21	--	.15	.80	.21	.49	--	.14	--
Significance	.748	.586	.283	.508	.833	--	.878	.424	.832	.622	--	.885	--

* p .05 before correction for multiple comparisons.

Note. ±Code to subtests: 1 = information; 2 = similarities; 3 = arithmetic; 4 = vocabulary; 5 = comprehension; 6 = digit span; 7 = picture completion; 8 = picture arrangement; 9 = block design; 10 = object assembly; 11 = coding A; 12 = coding B; and 13 = mazes.

Table 7

Standard Deviations of WISC IQ Scores, Program vs. Control by Project

Project		Mean Age	Mean s.d. Program	Mean s.d. Control	F Value	Significance	N Program	N Control
Gordon	FIQ	10-4	13.10	11.09	1.40	.420	70	20
	VIQ	10-4	12.95	10.82	1.43	.384	70	20
	PIQ	10-4	13.27	11.85	1.25	.595	70	20
Gray	FIQ	16-9	12.11	15.39	1.61	.233	34	18
	VIQ	16-9	9.59	16.03	2.79	.011*	34	18
	PIQ	16-9	16.76	15.31	1.20	.707	34	18
Levenstein	FIQ	9-9	10.98	9.70	1.28	.516	51	25
	VIQ	9-9	10.80	8.55	1.59	.215	51	25
	PIQ	9-9	13.07	12.68	1.06	.894	51	25
Miller	FIQ	12-8	11.38	14.11	1.54	.110	109	32
	VIQ	12-8	12.81	12.23	1.10	.792	109	32
	PIQ	12-8	12.53	15.33	1.50	.134	109	32
Palmer	FIQ	12-2	12.28	16.20	1.74	.050*	104	28
	VIQ	12-2	12.84	15.23	1.41	.228	104	28
	PIQ	12-2	13.63	13.09	1.08	.845	104	28
Weikart	FIQ	14-0	11.19	10.90	1.05	.846	54	56
	VIQ	14-0	9.74	10.45	1.15	.604	54	56
	PIQ	14-0	13.21	11.81	1.25	.411	54	56

* $p < .05$ before correction for multiple comparisons.

Table 8

Standard Deviations of WISC Subtest Scores, Treatment vs. Control by Project

Project	Subtest*												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Gordon													
Control	2.26	3.42	2.59	2.20	1.80	--	2.61	3.16	3.27	3.03	--	3.04	--
Program	2.53	2.47	2.70	2.74	2.77	--	2.84	3.52	3.07	2.43	--	2.62	--
F Value	1.26	1.91	1.08	1.56	2.36	--	1.18	1.24	1.13	1.54	--	1.35	--
Significance	.600	.059	.888	.290	.045*	--	.714	.631	.682	.204	--	.373	--
Gray													
Control	2.71	3.05	3.07	2.96	3.19	--	3.36	2.69	2.43	2.80	3.93	--	--
Program	1.60	2.19	2.41	2.18	1.91	--	2.92	3.68	2.97	3.74	3.11	--	--
F Value	2.87	1.94	1.62	1.85	2.78	--	1.32	1.87	1.49	1.79	1.60	--	--
Significance	.009	.101	.232	.129	.011*	--	.482	.172	.384	.205	.243	--	--
Levenstein													
Control	1.99	2.58	2.34	2.27	1.84	2.27	2.61	2.22	2.06	3.12	--	4.11	--
Program	2.41	2.45	2.42	2.71	2.49	2.53	2.09	2.64	3.09	3.03	--	3.03	--
F Value	1.47	1.11	1.07	1.42	1.82	1.24	1.56	1.42	2.25	1.06	--	1.84	--
Significance	.309	.733	.889	.350	.112	.581	.185	.357	.034*	.840	--	.069	--

Table 8 (Cont.)

Project	1	2	3	4	5	6	7	8	9	10	11	12	13
	Subtest												
Miller													
Control	2.06	2.94	2.28	2.30	3.34	3.37	2.77	3.29	3.41	3.03	--	2.97	3.05
Program	2.29	3.11	2.36	3.01	2.80	3.35	2.51	2.80	2.90	3.33	--	3.41	3.22
F Value	1.23	1.12	1.07	1.71	1.43	1.02	1.21	1.37	1.38	1.21	--	1.32	1.11
Significance	.518	.741	.850	.087	.187	.915	.461	.236	.232	.553	--	.374	.758
Palmer													
Control	3.93	3.18	2.58	3.25	3.35	2.24	2.18	2.68	2.75	3.15	--	3.27	3.25
Program	2.07	3.04	2.90	2.91	2.78	2.90	2.57	2.51	2.90	3.49	--	3.05	3.19
F Value	3.62	1.09	1.26	1.25	1.45	1.68	1.39	1.14	1.11	1.23	--	1.15	1.03
Significance	.000*	.724	.494	.422	.190	.122	.323	.617	.775	.554	--	.598	.865
Weikart													
Control	1.69	1.93	1.99	2.48	2.89	--	2.54	2.11	3.03	2.78	--	2.57	--
Program	1.82	2.08	2.12	1.95	2.05	--	2.60	2.74	2.60	2.51	--	3.88	--
F Value	1.16	1.16	1.13	1.61	1.99	--	1.05	1.69	1.36	1.22	--	2.28	--
Significance	.595	.588	.643	.084	.013*	--	.867	.055*	.259	.460	--	.003*	--

* p .05 before correction for multiple comparisons.

Note. Code to subtests: 1 = information; 2 = similarities; 3 = arithmetic; 4 = vocabulary; 5 = comprehension; 6 = digit span; 7 = picture completion; 8 = picture arrangement; 9 = block design; 10 = object assembly; 11 = coding A; 12 = coding B; and 13 = mazes.

Assignment to Special Education and Retention in Grade

The most exciting result from the previous Consortium analyses was the finding that low-income children who had participated in preschool programs were more able to meet the minimal requirements of their schools than were children in control groups. Treatment children were less likely to be retained in grade and less likely to be assigned to special education classes. These outcomes are important both to the child -- as concrete evidence of maintaining parity with peers and progressing satisfactorily -- and to society. For example, Weber, Foster, and Weikart (1978) conducted a benefit-cost analysis of the Perry Preschool program and concluded that (a) students who attended the preschool had higher projected lifetime earnings compared to students who had not attended preschool; and (b) economic benefits exceeded the project's costs because children who attended preschool did not require costly special education classes later in their school careers to the same extent that control children required such programs.

The initial Consortium findings on special education were based on data from five projects and those on retention in grade (grade failure) on seven projects. Those findings were reanalyzed to include additional data received since July, 1977. In addition, they were extended by investigating the effects of preschool experience on a new dependent variable -- underachievement. They were further elucidated by investigating whether preschool children maintained their superior school performance after controlling for background variables and IQ scores.

Methods

The variables used to assess school outcomes were assignment to special education (yes or no), retention in grade (yes or no) and underachievement (a composite of special education and/or retention in grade and/or dropping out of school before graduation from high school).[1] These data were transferred from school records onto the School Record Form by the field staff of the individual investigators.[2]

The use of these outcomes as measures of the effectiveness of early intervention programs has a major advantage over the use of IQ scores or achievement tests in that they are concrete indicators of whether a child has performed at an acceptable level within his/her educational institution. One drawback, however, is that special education placement and grade retention are affected by the policies of individual school districts and states. Therefore, in our analysis, comparisons are made within schools in the same district.

---

[1] Assignment to special education was coded if the School Record Form had been checked special education (unspecified), educable or trainable mentally retarded, learning disabled or emotionally disturbed. Classes for children with speech and hearing difficulties were not classified as special education. Retention in grade was coded if the student had been retained one or more times, with the exception of Palmer. The types of school failure classified as "retained" by Palmer were (a) administration of a lower level achievement test, relative to the child's age; (b) repetition of grade reported on Parent or Youth Interviews; (c) assignment to special education classes indicated on school records; and (d) achievement test scores more than 2 years behind grade level on current achievement test. This procedure was necessary due to the nature of Manhattan school records.

[2] See Appendix E for a copy of the School Record Form and Appendix A for attrition analyses of the School Record Form.

Contingency table analysis was used to test the main effects of early education programs on later school performance. In reporting results, data sets are divided into those which closely approximate experimental designs and those which are quasi-experimental. This procedure helps insure that the most rigorous test of the hypotheses will be limited to the data sets with random assignment. The unit of analysis was individuals within projects. As explained earlier, the results from each project were converted to a _z_ score, and the _z_ scores were summed across the projects. The combined significance level of statistical tests is reported separately for data sets which closely approximated experimental design and for experimental and quasi-experimental sets combined.

Results

Considering placement in special education classes, it is evident that the percentages of children so assigned varied considerably from project to project. This may be due to policy differences at the state and/or district levels or to initial differences between the samples. However, since treatment and control children are only compared within projects for these analyses, this variation presents no difficulties.

As Table 9 shows, a high percentage of control children had been assigned to special education classes -- a median of 28.6% across the six programs. The percentage of treatment children so assigned was in most cases substantially smaller -- a median of 13.8%.

Another way of representing this treatment-control difference is shown in the Percent Reduction column of Table 9. Here the percentage of control children in special education serves as a baseline to indicate the likely magnitude of special education assignments for low-income children with no preschool. By subtracting the percentage of treatment children actually assigned to special education, we arrive at a figure which indicates how much preschool programs can be expected to reduce later placement in special education classes. Calculated in this way, four programs substantially reduced such placements. The median of the four data sets was a percent reduction of 61.0.

As Table 9 indicates, the treatment-control differences were significant in four projects; 2 x 2 chi-square tests yielded two-tailed p values ranging from .0044 to .06.[1] Pooling the results from all six projects resulted in a highly significant p value of .0004. This was a robust finding. When the data set with the most significant p value (Gray) was omitted, the pooled p value remained highly significant ($p = .0098$). Thus, we can safely conclude that preschool intervention programs significantly reduced placements in special education programs.

---

[1] The Miller and Beller projects did not find differences in favor of treatment children. For Miller this result was probably due to the initial differences between Miller's treatment and control groups, differences which favored the control group. The Philadelphia schools in Beller's project either rarely or never utilized assignment to special education, or did not record the information on school records. The percentages involve the following actual numbers: one control child and two treatment children assigned to special education over a period of 12 years time.

The second measure of actual school performance is the percentage of children who have been retained in grade at least once during their school careers.[1] Across the eight projects represented in this analysis, a median of 30.5% of control children were retained compared to a median of 25.4% of treatment children. The median reduction in grade failure across all projects was 23.1%.

The grade-retention variable yields treatment-control differences similar in pattern to, but less striking than, those found for placement in special education. The data are presented in Table 10. All projects except Miller's reported that more control children than treatment children were retained in grade. Only Palmer's results were statistically significant with Weikart's results marginally significant. When the results of the four studies with the most rigorous designs were pooled, the average preschool intervention program significantly reduced retention in grade ($p = .0042$, two-tailed). Deleting the data set with the best p value (Palmer), a p of .0872 (two-tailed) resulted. Pooling across all eight projects resulted in a p value of .0184. However, when the strongest data set (Palmer) was omitted, the result was not significant ($p = .1416$, two-tailed).

Finally, a composite variable was created. Labelled "under-achievement," it includes students who were assigned to special education classes and/or retained in grade and/or dropped out of school.[1] As shown in Table 11, the difference between the percent of treatment and control children classified as underachievers was significant ($p = .0002$, two-tailed) and robust ($p = .0368$ after deleting the two most significant results -- Palmer and Weikart in this case).

---

[1] There is considerable variability across projects, ranging from no children retained (Miller control group) to 68.6% retained (Gray control group). This variability reflects known variation across school districts in policies pertaining to retention.

In summary, the new data analyses confirmed our earlier finding: children who participated in preschool intervention programs were more likely than control children to meet at least the minimal standards of their schools. This is especially true in the case of assignment to special education classes, where the effects of preschool were highly significant, robust, and large. Results for retention-in-grade, while statistically significant across the projects, were only moderately robust. Analysis of the composite variable of underachievement resulted in significant and robust treatment/control differences.

In retrospect, we feel that the weaker findings for retention-in-grade may perhaps be explained as follows. First, retention in grade is less likely to differentiate between treatment and control children because of widespread use of social promotion. Also, many more control than treatment children are assigned to special education classes. In our analysis, those in regular classrooms who failed to pass were coded as retained in grade.[1] Children in special education classes who remained in special education classes were not coded as retained in grade. Since significantly more control children were assigned to special education, this raises the possibility that some of the worst control students (in the research design sense) have been differentially removed from the pool of students whom it was possible to retain. This is analogous to an attrition of the control subjects with lowest IQs: subsequent analyses will make the effect of preschool seem smaller than it really was.

---

[1] Beller and Gray provided data on subjects who were old enough to drop out of school. Only two Beller cases and two Gray cases dropped out of school without ever repeating a grade or attending special classes.

It is important to give some thought to the meaning of these findings. Bombarded by many hypothesis tests and results, the reader might be inclined to merely note the significant result and pass on. Instead, let us pause to briefly outline the social impact of this particular result.

Many studies confirm that over the last 15 years teachers assigned students to special education classes for a variety of reasons (cf. Milofsky, 1974). Assignment procedures varied so widely -- from building to building as well as across school districts and across states -- as to appear arbitrary and capricious.[2] IQ scores were widely used to designate children as mental retardates in need of special education (with a resultant over-representation of minority children in such classrooms). In addition, special education classes were widely used as a "dumping ground" for disruptive children, motivated, some suggest, by a wish to get rid of "undesirables" rather than to meet their special needs.

---

[1] This includes students who at one time had been in special education classes but had subsequently been moved back into regular classrooms.

[2] The Education for All Handicapped Children Act was passed in 1975 in partial response to such practices.

It is fair to assume that the mere assignment to such classes affects children. They are labelled in their own eyes and the eyes of others. Labels such as "emotionally disturbed" or "mildly retarded" have a life of their own, remaining on children's records for years and potentially affecting each new teacher's expectations for and treatment of such children.

While not as radical as assignment to special classes, retention in grade may also be a blow to a child's feelings of self-esteem and self-worth and an occasion for teasing and ridicule from peers. Furthermore there are some indications that grade retention is no more beneficial than grade promotion, especially after the primary grades (Kraus, 1973).

The value of staying in rather than dropping out of high school is self-evident. The individual high school graduate has broader choices and life-chances compared to the dropout. The society also gains by having members who are literate, informed and employable.

Thus, by avoiding placement in special education classes, retention in grade, and dropping out of school, children are by-passing swamps where many poor, minority children flounder: the weight of an official label, a detour into a different system and difficulties getting back "on track", the feelings of personal failure and worthlessness.

Some Tests of Alternative Hypotheses

The analyses just reported might be criticized on various grounds. This section considers several potential criticisms and reports reanalyses designed to evaluate their reliability.

First, it is possible to conjecture that all these findings on preschool effectiveness were caused by nonrandom differences between treatment and control groups when perfect randomization did not occur. In order to investigate this possibility, the data were reanalyzed using standard regression analysis. This procedure allowed simultaneous consideration of the effects of preschool, the child's sex, ethnicity, and cognitive ability (as measured by initial IQ score),[1] family size (number of siblings), family structure (father absent or present), and the level of education attained by the child's mother.[2] The effectiveness of preschool on school outcomes, independent of the effects of the other independent variables, can then be tested by testing the statistical significance of the unstandardized regression coefficient for preschool vs. control.[1]

---

[1] Stanford-Binet IQ scores except for Levenstein's project, for which Peabody Picture Vocabulary Test scores were used.

[2] A somewhat smaller number of children are included in the regression analyses due to missing data on one or more independent variables for about 50 students. (Attrition analyses [Appendix A] found no differential attrition.) The Gordon and Miller data sets were not used in this analysis because they had no pretest IQ scores.

As Table 12 shows, the results were highly significant: early education positively affected school performance independent of the effects of sex, ethnicity, family background and initial intelligence. With assignment to special education as the dependent variable, the result was significant ($p = .0001$) for both the two experimental data sets and the combined experimental and quasi-experimental data sets ($p = .0001$). It was also robust; when the strongest data set (Gray) was deleted, the result remained significant ($p = .0226$). The results for underachievement were similarly significant ($p = .0022$) and robust ($p = .026$ after deleting the strongest data set -- Weikart). For retention in grade, the results were statistically significant ($p = .0235$) but only moderately robust ($p = .105$, two-tailed, after deleting the strongest data set -- Palmer).

Since the Miller and Gordon projects did not have pretest IQ data, they were not included in the analysis above. However, more regression analyses which included these two groups were performed in order to test the effects of the various background variables (minus pretest IQ) on school outcomes. As Table 13 shows, the outcome was similar. Independent of sex, ethnicity and family background measures, preschool positively affected school outcomes. Again, the results for placement in special education classes and for the composite underachiever variable were both significant and robust while results for retention in grade were statistically significant[1] but only moderately robust (p = .063, two-tailed).

---

[1] The computer program yielded a t for testing effectiveness of preschool. This t was converted to a standard normal deviate by the formula $|z| = df*ln(t^2/df+1)$ with sign (z) = sign (t). See Darlington, 1978(b).

On the basis of these two sets of analyses, we can safely conclude that children benefit from preschool programs -- in being more likely to meet the minimal requirements of later schooling -- and that this finding was not due to initial treatment/control differences in sex, ethnicity, early family background and early intelligence level.

A second potential criticism is as follows. Suppose early education programs have a short-term effect on test-taking ability which appears as a temporarily inflated IQ test score. If a child's IQ score at first grade is in his folder, it might influence teachers to keep the "brighter" children out of special education classes.[2] If this hypothesis were true, then removing the influence of the IQ score at 6 years old would remove any association between attending preschool and placement in special education.

This hypothesis was tested using a standard regression analysis with the following independent variables: preschool, IQ score at age 6, mother's education, father presence, number of siblings, sex of child, and ethnicity of child. This is an extremely rigorous test of the effects of preschool because, to the extent that IQ tests are a broad-scale measure of cognitive ability, partialling out the effect of IQ score at age 6 means removing a large portion of the cognitive effects of preschool programs (It also removes the effect of cognitive ability which is necessary to perform acceptably in school). In other words, we deliberately committed the "mistake" of partialling out a variable affected by the independent variable of interest; this makes the test more conservative. As Table 14 shows, preschool still affected special education placement independently of the effects of sex, ethnicity, family background variables and IQ score at age 6 ($p = .0192$, two-tailed). When the data set with the strongest p value was deleted (Levenstein), the significance level drops to $p = .2380$, two-tailed.

---

[1] Special education, p = .0028; grade retention, p = .0387; underachievement, p = .0031

[2] In some school districts it is stated policy that children with IQ scores above 80 (for example) may not be assigned to special education.

Since this is such a conservative test, it is worthwhile to consider the results from the projects which most closely approximated true experimental designs. As the table shows, for those four projects, preschool affected special education placement independently of its effect on IQ scores at age 6 ($p = .0298$, two-tailed). When the strongest data set (Gray) was eliminated, the p value dropped to .1590, two-tailed, or .0795, one-tailed.

When IQ score at age 6 was partialled out of the equation for grade retention, the results were not significant. They were marginally significant for the underachievement variable ($p = .0728$, two-tailed for the approximately experimental data sets and $p = .0932$ for all data sets).

Partialling out IQ score at age 6 in predicting assignment to special education served to test the labeling hypothesis that teachers identified treatment children as brighter and so were less likely to recommend them for placement in special classes. The results disproved the hypothesis; preschool experience affected special education placement apart from IQ score at age 6. In effect, this means that preschool must have affected the children beyond the purely cognitive influences of teaching concepts and skills, yet in ways which were related to school performance.[1]

In a sense, partialling out IQ score at age 6 has a different meaning when applied to retention in grade. Presumably, teachers primarily retain children in grade not because of differential labeling, but because by some objective criteria the children failed to master the material. Partialling out IQ score at age 6 in effect removes the influence of cognitive ability necessary to master school material. Thus we should not really expect a significant effect of preschool on grade retention when IQ at age 6 is controlled.

A third potential challenge to the finding in this section might concern the use of the individual child as the unit of analysis. As outlined earlier, this challenge might stress the fact that an especially capable teacher or other chance event could affect an entire classroom of children, so that the individual children in the classroom are not really independent of each other in the statistical sense. To answer this, we might use the classroom as the unit of analysis, bringing the N down from about 2,000 to about 200. But classrooms might not be independent of each other. The hypothetical capable teacher may have taught in several different classes over several years, or the supervisor of, for example, a Montessori program may have been particularly adept at selecting excellent teachers. In addition, there is practical problem that not all investigators could supply records on the exact classrooms in which each child had been placed.

---

[1] This topic is taken up later in this report when we consider preschool's effect on the children's achievement motivation, self-evaluation, and sociability, and on their parent's aspirations.

The next larger unit of aggregation is the program within a site -- for instance, Montessori, Bereiter-Engelmann, or DARCEE.[1] Upon consideration, this seemed to be the most appropriate unit of analysis.

For performing an analysis at the program level, we had data on special education placements for 21 programs at 7 different sites. These 21 programs are shown in Table 28 at the end of the chapter on program characteristics. Under the "Adjusted Difference" column on Table 28 is a number representing the effectiveness of that program in reducing special education placement. Effectiveness was defined as the proportion of experimental-group children not placed in special education classes, minus the comparable proportion for control-group classes, after analysis of covariance was used to adjust for any differences between experimental and control groups on pretest IQ scores. (The five Karnes programs and the five Miller programs were not used in this analysis. Karnes had no control group and Miller's control group was found post hoc to be different from the treatment group in ways favoring the controls. This left 11 programs at five sites in the analysis.)

---

[1] The next larger unit of analysis after programs is the site itself, which brings our total N down to only 12, or fewer for most analyses, since not all investigators supplied us with data for all analyses. This is clearly too small, and it also seems meaningless to lump completely different programs just because they occurred at the same site.

Tabel 28 shows that all 11 programs had positive effectiveness; that is, with remarkable consistency these programs appeared to be at least somewhat effective in reducing special education placements. This result was statistically significant by the sign test: $p < .001$, two-tailed. The median program was Gray's two-year program. In this program, .298 of the control group children were placed in special education classes compared to only .053 of the experimental-group children. (The difference of .241 between these two proportions became .250 when adjusted for the slight difference between the two groups in mean pretest IQ.)

We may summarize our analyses of the effect of preschool experience on later school outcomes as follows. These analyses have been performed with both individuals and subgroups of the projects as units of analysis. Some analyses have controlled for pretest IQ and family background variables. Some have even controlled for posttest IQ. All of these analyses have yielded the same basic conclusion: preschool makes a positive contribution to low-income children's later school outcomes.

Table 9

Percentage of Subjects Placed in Special Education Classes, Treatment vs. Control

Data Sets (n)	Treatment	Control	Percent Reduction*	Chi-Square**	p (2-tailed)	Pooled z	Pooled p (2-tailed)
Closely approximating experimental design							
Gordon (82)	23.2%	53.8%	56.9%	5.10	.0244		
Gray (53)	2.8	29.4	90.5	8.16	.0044		
Weikart (123)	13.8	27.7	50.2	3.55	.0602		
Median (258)	13.8	27.7	56.9			4.04##	<.0001
Quasi-experimental							
Beller (66)	5.7	3.2	----	.23	(.6315)#		
Levenstein (125)	13.7	39.1	65.0	8.07	.0046		
Miller (109)	20.4	12.5	----	.55	(.4654)		
Median for all Data Sets (558)	13.8	28.6	61.0#			3.52	.0004

Note: Data were collected when most of the children were in the following grades: Gordon, 5th grade; Gray, 12th grade; Weikart, 4th grade; Beller, 12th grade; Levenstein, 3rd grade; and Miller, 7th grade.

* Percent reduction = [(% control - % treatment/ % control].

** Without Yates correction. See Camilli & Hopkins, 1978.

# Figures in parentheses are in the reverse direction; that is, treatment was associated with an increase in the rate of assignment to special education. Miller and Beller results are excluded from median percent reduction.

## Pooled $z = \Sigma z_i / \sqrt{k_i}$ where z_i = z score from sample i and k is number of data sets.

100

Wat

Table 10

Percentage of Subjects Retained in Grade, Treatment vs. Control

Data Sets (n)	Treatment	Control	Percent Reduction*	Chi-square**	p (2-tailed)	Pooled z	Pooled p (2-tailed)
Closely approximating experimental design							
Gordon (65)	27.6%	28.6%	3.5%	0.00	.9563		
Gray (50)	52.9	68.8	23.1	1.12	.2907		
Palmer (221)	24.1	44.7	46.1	7.66	.0056		
Weikart (97)	4.0	14.9	73.2	3.42	.0646		
Median (433)	25.9	36.7	34.6			2.87##	.0042
Quasi-experimental							
Beller (66)	42.9	51.6	16.9	0.51	.4751		
Levenstein (109)	12.9	18.8	31.4	0.39	.5307		
Miller (106)	7.8	0.0	----	1.33	(.2484)#		
Zigler (144)	26.6	32.3	17.6	0.57	.4519		
Median for All Data Sets (858)	25.4	30.5	23.1#			2.36	.0184

Note: Data were collected when most of the children were in the following grades: Gordon, 5th grade; Gray, 12th grade; Palmer, 7th grade; Weikart, 4th grade; Beller, 12th grade; Levenstein, 3rd grade; Miller, 7th grade; and Zigler, 7th & 8th grade.

* Percent reduction = [(% control - % treatment)/% control].

** Without Yates correction. See Camill[illegible] & Hopkins, 1978.

# Figures in parentheses are in the reverse direction; that is, treatment was associated with an increase in the rate of retainment in grade. Miller results are excluded from median percent reduction.

## Pooled $z = \frac{\sum z_i}{\sqrt{k_i}}$ = z score from sample i and k is number of data sets.

Table 11

Percent of Underachieving Students (Placed in Special Education Classes, and/or Retained in Grade and/or Dropped Out of School), Treatment vs. Control

Data set (n)	Treatment	Control	% Reduction*	Chi-square**	p (2-tailed)	Pooled z	Pooled p (2-tailed)
Closely approximating experimental design							
Gordon (82)	39.1	61.5	36.0	2.25	.1340		
Gray (55)	55.6	73.7	24.6	1.73	.1882		
Palmer (221)	24.1	44.7	46.1	7.66	.0056		
Weikart (123)	17.2	38.5	55.3	6.78	.0092		
Median (481)	31.6	53.1	41.1			4.09##	<.0001
Quasi-experimental							
Beller (69)	48.6	53.1	8.5	0.14	.7107		
Levenstein (127)	22.1	43.5	49.2	4.47	.0345		
Miller (125)	20.6	11.1	----	0.89	(.3455)#		
Zigler (144)	26.6	32.3	17.6	0.57	.4502		
Median for All Data Sets (920)	25.4	44.1	36.4			3.71	.0002

Note: Data were collected when most of the children were in the following grades: Gordon, 5th grade; Gray, 12th grade; Palmer, 7th grade; Weikart, 4th grade; Beller, 12th grade; Levenstein, 3rd grade; Miller, 7th grade; and Zigler, 7th & 8th grade.

* Percent reduction = [(% control - % treatment)/% control].

** Without Yates correction. See Camilli & Hopkins, 1978.

# Figures in parentheses are in the reverse direction; that is, treatment was associated with an increase in the rate of underachievement. Miller results are excluded from median percent reduction.

## Pooled $z = \Sigma z_i / \sqrt{k_i}$ where z_i = z score from sample i and k is number of data sets.

Table 12

The Effect of Early Education on School Outcomes When Background Variables and Pretest IQ Score Are Controlled

Data Set	n	Preschool Coeff.	F	Signif. Level (2 tailed)	Pooled z score	Pooled p (2 tailed)
			Special Ed Placement			
Approx. Experimental Design						
Gray	47	.396[a]	17.551[b]	.0002		
Weikart	123	.119	2.919	.0909		
Experimental Total	170				3.8651[d]	.0001
Quasi-Experimental						
Beller	52	.005	.005	.9443		
Levenstein	119	.206	4.946	.0285		
All Data Sets	341				3.863	.0001
			Grade Retention			
Approx. Experimental Design						
Gray	44	.110	.460	.5046		
Weikart	97	.103	2.890	.0935		
Palmer	132	.161	3.434	.0667		
Experimental Total	273				2.4122	.0159
Quasi-Experimental						
Beller	52	.026	.038	.8472		
Levenstein	104	.062	.486	.4885		
All Data Sets	429				2.2644	.0235
			Underachievement (Special Ed/Retention/Dropout)			
Approx. Experimental Design						
Gray	48	.137	.861	.3617		
Palmer	132	.161	3.434	.0667		
Weikart	123	.191	6.033	.0157		
Experimental Total	303				2.9794	.0029
Quasi-Experimental						
Beller	55	-.018	.019	(.8915)[c]		
Levenstein	121	.189	3.449	.0665		
All Data Sets	479				3.0676	.0022

Note. Equation: ACH = MED + FP + SIBS + SEX + ETH + IQPRE + PC.
Legend: ACH = Special education placement, in-grade retention, composite under-achievement; MED = mother's education; FP = father presence; SIBS = number of siblings; ETH = ethnicity; IQPRE = pretest IQ score; IQ6 = IQ score at 6 years old; PC = preschool vs. control. IQ scores = Stanford Binet (except PPVT for Levenstein).

[a] Unstandardized regression coefficient for preschool vs. control.

[b] F test for significance of preschool coefficient, simultaneous (standard regression) method. Figures in table are as if preschool were added after all other variables.

[c] Figures in parentheses are in reverse direction, i.e. preschool had a negative association with school outcomes.

[d] Pooled $z = \Sigma z_i / \sqrt{k}$.

Table 13

The Effect of Early Education on School Outcomes When Background Variables Alone Are Controlled

Data Set	n	Pre-School Coeff.	F	Signif. Level (2 tailed)	Pooled z score	Pooled p (2 tailed)
Special Ed Placement						
Approx. Experimental Design						
Gordon	64	.174 [a]	.937 [b]	.3391		
Gray	51	.288	9.745	.0033		
Weikart	123	.132	3.485	.0650		
Experimental Total	238				3.3150 [d]	.0009
Quasi-Experimental						
Beller	53	-.011	.027	(.8709) [c]		
Levenstein	123	.238	6.953	.0097		
Miller	106	-.100	.736	(.3942)		
All Data Sets	520				2.9864	.0028
Grade Retention						
Approx. Experimental Design						
Gordon	53	.075	.112	.7406		
Gray	48	.178	1.303	.2628		
Palmer	219	.199	7.227	.0078		
Weikart	97	.108	3.222	.0767		
Experimental Total	417				2.9403	.0033
Quasi-Experimental						
Beller	53	.023	.029	.8663		
Levenstein	107	.061	.445	.5073		
Miller	101	-.077	1.286	(.2142)		
All Data Sets	678				2.0674	.0387
Underachievement (Special Ed/Retention/Dropout)						
Approx. Experimental Design						
Gordon	64	.097	.219	.6429		
Gray	52	.208	2.171	.1495		
Palmer	219	.199	7.227	.0078		
Weikart	123	.205	6.772	.0106		
Experimental Total	458				3.5600	.0004
Quasi-Experimental						
Beller	56	-.029	.048	(.8284)		
Levenstein	125	.216	4.441	.0376		
Miller	120	-.123	1.347	(.2493)		
All Data Sets	759				2.9595	.0031

Note. Equation: ACH = MED + FP + SIBS + SEX + ETH + PC. For legend see Table 12.

[a] Unstandardized regression coefficient for preschool vs. control.

[b] *F* test for significance of preschool coefficient, simultaneous (standard regression) method. Figures in table are as if preschool were added after all other variables.

[c] Figures in parentheses are in reverse direction, i.e. preschool had a negative association with school outcomes.

[d] Pooled z = $\Sigma z_i / \sqrt{k}$.

Table 14

The Effect of Early Education on School Outcomes When Background Variables and IQ Score at Age 6 Are Controlled

Data Set	n	Pre-School Coeff.	F	Signif. Level (2 tailed)	Pooled z score[d]	Pooled p (2 tailed)
			Special Ed Placement			
Approx. Experimental Design						
Gordon	62	.198[a]	1.059[b]	.3101		
Gray	50	.183	3.308	.0775		
Weikart	120	.069	.981	.3251		
Experimental Total	232				2.1734[d]	.0298
Quasi-Experimental						
Beller	50	-.022	.093	(.7633)[c]		
Levenstein	116	.335	10.002	.0021		
Miller	106	-.094	.660	(.4197)		
Add Data Sets	504				2.3416	.0192
			Grade Retention			
Approx. Experimental Design						
Gordon	51	-.014	.004	(.9502)		
Gray	46	-.070	.179	(.6765)		
Palmer	195	.148	3.348	.0692		
Weikart	94	.080	1.616	.2083		
Experimental Total	386				1.2977	.1944
Quasi-Experimental						
Beller	50	-.063	.250	(.6217)		
Levenstein	101	.140	1.433	.2355		
Miller	101	-.064	.911	(.3436)		
All Data Sets	638				.8848	.3763
			Underachievement (Special Ed/Retention/Dropout)			
Approx. Experimental Design						
Gordon	62	.085	.161	.6911		
Gray	50	-.043	.076	(.7853)		
Palmer	195	.148	3.348	.0692		
Weikart	120	.126	2.768	.0997		
Experimental Total	427				1.7943	.0728
Quasi-Experimental						
Beller	53	-.085	.484	(.4925)		
Levenstein	118	.309	6.519	.0122		
Miller	120	-.102	.947	(.3336)		
All Data Sets	718				1.6785	.0932

Note. Equation: ACH = IQ6 + MED + FP + SIBS + SEX + ETH + PC. For legend see Table 12.

[a] Unstandardized regression coefficient for preschool vs. control.

[b] F test for significance of preschool coefficient, simultaneous (standard regression) method. Figures in table are as if preschool were added after all other variables.

[c] Figures in parentheses are in reverse direction, i.e. preschool had a negative association with school outcomes.

[d] Pooled $z = \Sigma z_i / \sqrt{k}$.

Achievement Tests

A primary goal of most infant and preschool intervention programs was to improve children's later school performance and thus circumvent the familiar pattern of low-income children falling farther behind in school with each passing year. Avoiding retention in grade and placement in special education classes has been used as one indicator of meeting the minimal requirements of school. Children's performance on standardized achievement tests is another basic indicator of whether preschool programs were successful in improving children's educability.

Many of the individual projects which comprise the Consortium on Longitudinal Studies had published the results of their comparisons of treatment and control children's scores on achievement tests. Two projects reported achievement test scores over time, and their results are apparently contradictory. Gray's treatment children initially scored higher than controls, but their advantage disappeared over time. The children were administered the Metropolitan Achievement Tests (MAT) during the springs of first grade, second grade and fourth grade. Treatment children were significantly superior on three out of four subtests in first grade and on two out of five subtests in second grade. By the fourth grade, treatment children still scored above control children, but the differences were not statistically significant (Gray, Klaus & Ramsey, 1978). Achievement test scores showed something of a "sleeper effect" in Weikart's project. Children were given the California Achievement Test (CAT) in grades one through eight. In this case, treatment children's scores were virtually the same as the scores of control children at the end of first grade. With each succeeding year the differences in favor of children with preschool became larger; in the eighth grade, children with preschool had significantly higher scores on all three (reading,[1] language and arithmetic) areas covered by the CAT test (Schweinhart & Weikart, 1978).

Results from other projects were varied. Levenstein (1978) reported significant differences in favor of the treatment children on both the reading and arithmetic subtests of the Wide Range Achievement Test administered at the end of the third grade. Palmer (1977) found that fifth grade treatment children scored significantly better than controls on arithmetic achievement. In the seventh grade treatment children scored significantly higher than controls on reading achievement, using the California Test of Basic Skills (Palmer, et al., 1978).[1] Miller (1977), however, reported that her treatment and control subjects did not differ on Stanford Achievement Test scores for either reading or math at the end of sixth grade.[2]

---

[1] Both Weikart and Palmer found significant treatment/control differences on reading vocabulary but not reading comprehension subscores.

[2] It is important to note once again that Miller's control group was initially somewhat different from the treatment group, with differences favoring the controls.

What, then, can we conclude about the effect of preschool on achievement test scores? Once again, as in other analyses, we attempted to pull together the disparate findings by testing the null hypothesis that there is no "average" effect of preschool on children's achievement test performance.

Methods

Since achievement tests were ordinarily administered by the public school systems, most of the investigators were unable to control the precise achievement tests given to their subjects.[1] The school systems also controlled the grades in which the tests were given. Nevertheless, most investigators were able to find at least one grade in which most of their subjects had been given the same test, and to supply the data to us in a form which we were able to analyze. In particular, there were useful amounts of achievement test information in the fourth grade for six investigators -- Beller, Gordon, Levenstein, Miller, Palmer, and Weikart.[2] This included both mathematics and reading tests for all six investigators. For Palmer, however, the mathematics test information was most usable in the fifth grade rather than the fourth.

Several Consortium members supplied data for two types of reading tests -- scores on reading comprehension and scores on broader language-ability tests that included other language skills such as vocabulary. Since there were more data on broad-gauge tests than on the narrower tests, we decided to limit the present analysis to the broader tests.

---

[1] The Weikart project administered achievement tests to their own subjects.

[2] Due to time constraints, we were unable to include the Gray fourth grade achievement test data in these analyses.

Of the subjects used in this analysis, all from Beller's project took the 1964 edition of the Iowa Test of Basic Skills; all from Gordon's took the 1965 edition of the Metropolitan Achievement Test; all from Levenstein's took the 1965 edition of the Wide Range Achievement Test; all from Miller's took the California Test of Basic Skills; and all from Weikart's took the 1957 edition of the California Achievement Test. Some of Palmer's subjects took the 1970 edition of the Metropolitan Test, and some took the 1972 edition of the Stanford Achievement Test. For the Palmer project only, a dummy variable was included in the regression equation measuring which of these two tests was taken. This meant that data for the two tests were in effect analyzed separately and the results then combined.

Regression analysis was used separately for each investigator and for each test -- math or reading. Since this analysis involved six investigators, this made a total of 12 regression equations. The independent variable was preschool attendance. Age was included as a covariate or control variable partly to correct for any bias introduced by the fact that some children might score better on fourth grade achievement tests because they were the age of typical fifth graders but had been retained in grade a year. In addition, the child's sex and pretest IQ (available only for

Beller, Miller[1], and Palmer) were entered into the regression as covariates.

For each regression, the computer program yielded a t for testing the effectiveness of preschool. This t was converted to a standard normal deviate by the formula $|z| = df*ln(t^2/df+1)+1)$, with sign (z) = sign (t) (from Darlington, 1978b). The six z scores thus obtained for any one type of test (math or reading) were then combined in two different ways, depending on the unit of analysis. When the individual child was the unit of analysis, the scores were combined by the formula $z = \Sigma z/\sqrt{k}$ mentioned in the section on methods. When the project was the unit of analysis, the mean and standard deviation of the six scores were computed just as if the scores were six raw scores on some measure. An ordinary t test was then used to test the null hypothesis that the mean of the five scores was 0.

Results and Conclusions

The results of these analyses are shown in Table 15. Treatment children showed a significant advantage over control children in mathematics achievement, while the verbal achievement scores showed a suggestive trend in the same direction. The difference in mathematics scores was highly significant regardless of whether the unit of analysis was the individual child (p = .008, two-tailed; p = .0016, one-tailed) or the project (p = .007, two-tailed; p = .0014, one-tailed).

---

[1] The reader will recall that Miller administered IQ tests to treatment and control children after treatment children had been enrolled in preschool for 6 to 8 weeks. Thus, the pretest IQ measure was not statistically independent of the effects of the treatment. This artifact for this particular analysis would favor control rather than treatment children. In general, covariance analysis undercorrects for such biases, and so the analysis still slightly favors the controls.

For reading ability, the child-level and project-level p values were .14 and .30 (two-tailed) or .07 and .15 (one-tailed) respectively. The former figure could be considered at least a suggestive trend.

It should be pointed out that the significance levels observed with mathematics were small enough to survive correction for the fact that they were the better of two results -- math and reading. We thus conclude that on the average preschool does appear to have some positive effect on achievement test scores.

Although significant, are the effects under discussion large enough to be of practical interest? The dependent variables in our analyses were raw scores or simple monotonic transformations of raw scores. Thus we were unable to compare the sizes of the effects across projects because different projects used different achievement tests. However, some rough idea of the size effect we are considering can be gained as follows. On the mathematics test, if we rank the six projects in the order of the significance of their differences, one of the median projects is Beller's. In the Beller project, the difference between treatment and control groups, controlling for age, sex, and pretest IQ, was .52 grade-equivalents. This figure gives some idea of the size effects we are considering. Even if estimation error has affected this quantity somewhat and the true difference is only half as large as .52, the difference is still large enough to be of considerable practical interest.

Table 15

Fourth Grade Treatment-Control Differences on Achievement Tests

Mathematics

Project	*N*	*t*	*df*	*z*
Beller	50	1.620	45	1.580
Gordon	41	1.[illegible]1	36	1.804
Levenstein	45	[illegible]0	40	.296
Miller	82	.894	77	.886
Palmer*	89	1.437	84	1.420
Weikart	95	1.233	90	1.221

Analysis with child as unit of analysis:

Sum of 6 z's = 7.207; $z = 7.207/\sqrt{6} = 2.942$; $p = .008$, two-tailed

Analysis with project as unit of analysis:

Mean of 6 z's = 1.201; Standard deviation of 6 z's = .543

$t = 1.201/(.543/\sqrt{6}) = 5.419$; df = 5; $p = .007$, two-tailed

Reading

Project	*N*	*t*	*df*	*z*
Beller	50	.804	45	.793
Gordon	41	.613	36	.608
Levenstein	45	2.170	40	2.085
Miller	82	.000	77	.000
Palmer	89	-1.442	84	-1.424
Weikart	95	1.542	90	1.523

Analysis with child as unit of analysis:

Sum of 6 z's = 3.585; $z = 3.585/\sqrt{6} = 1.464$; $p = .14$, two-tailed

Analysis with project as unit of analysis:

Mean of 6 z's = .598; Standard deviation of 5 z's = 1.229

$t = .598/(1.229/\sqrt{6}) = 1.191$; df = 5; $p = .30$, two-tailed

* Fifth grade data

NON-COGNITIVE EFFECTS OF PRESCHOOL

Attitudes and Values

We turn now to an investigation of non-cognitive outcomes of preschool intervention programs. Two related interests have guided this search. First is the fact that many intervention programs, including those comprising the Consortium, specifically set non-cognitive goals such as increasing children's self-esteem (Beller, Deutsch, Gray), enhancing social and emotional development (Gordon, Karnes, Miller) and influencing attitudes related to school success (Gray). As mentioned previously, Head Start goals also include the stimulation of social development. Clearly, it is of interest to know whether preschools succeeded in meeting these goals.

Additionally, we approach non-cognitive outcomes in the context of strong and robust evidence that preschool positively affected children's school performance and that it affected other than cognitive attributes -- i.e., effects existed even after partialling out the effects of preschool on IQ. There can be no question of the outcome. But by what process did it come about? It seems unlikely that one or at most 2 years of preschool experience could protect children against future school failure for the succeeding 10 to 15 years. Perhaps analysis of non-cognitive outcomes will provide clues to the nature of the intervening processes.

In most cases, the early intervention programs inaugurated in the 1960s made assumptions, implicitly or explicitly, about intervening processes which would explain the school failure of low-income children. Low-income children were described as in need of enrichment. Because of the nature of their past experience, the children lacked school-relevant skills and concepts, motivations, and goals for the future. Consequently, they were at high risk of school failure (see for example, Bloom, Davis, & Hess, 1965). It was reasoned that by intervening early in the child's life, s/he could be provided with learning experiences relevant to school and the parents could be taught how to support and teach him/her more effectively.

The preceding description contained the following simple model for social change. Preschool programs teach children concrete skills and concepts. But skills and concepts must be built upon over the years, so children must also be motivated to continue to learn and achieve in school. They must believe that school is important and possess enough self-confidence to exert the necessary effort. The preschool experience should also affect parents so that they may support their children's efforts. With the backing of new abilities, motivations, values and parental support, children should be better able to compete with their middle-class peers.

This simple model may be tested to some degree and non-cognitive outcomes assessed by using data from the Youth Interviews and the Parent Interviews collected by the Consortium in the 1976-77 follow-up study.[1] The Youth Interview covered educational expectations, occupational aspirations, attitudes toward school, current employment status, leisure time activities and interests, social interaction with family, peers and the larger community, and attitudes toward oneself and others. The Parent Interview covered such topics as household composition, socio-economic status, parental aspirations for and evaluations of their child, information on the child's medical history, the parent's current relationship with the child, and parental assessment of the intervention program.

---

[1] For a copy of the two interviews, see Appendix F.

Four content areas were explored in the current analyses: (a) mothers' aspirations for their children, (b) children's achievement-orientation, (c) children's self-evaluations, and (d) children's social relationships and social participation.

Maternal aspirations for their children have been included because so many intervention programs were explicitly aimed at fostering change in parents as well as children. Further, our model stipulated that persistent treatment effects of preschool intervention programs are partially mediated through the family. One possible route, for example, may be that preschool programs initially improve the child's school performance, thus affecting the mother's expectancies. This change in maternal attitude (and, presumably, behavior) could in turn constitute a "treatment" lasting well beyond the intervention program itself.

Achievement-orientation includes achievement motivation[1] and those values, attitudes, norms and goals which seem important for success in school and later jobs (cf. Kahl, 1965). In general, empirical evidence supports the conclusion that achievement orientation in its different forms can play an appreciable independent role in determining academic success (Spenner & Featherman, 1977). Our "model" predicted that treatment children would show more evidence of achievement orientation than controls.

Self-evaluation (or self-concept) has also been identified as an important factor contributing to academic success (e.g., Raizen, et al., 1974). Anderson and Evans (1976) posited a model similar to the one presented here in which self-concept is both a cause of achievement -- through its effect on achievement-orientation -- and an effect of actual achievement.

Sociability and social participation are not so clearly related to academic success as the previous three constructs. Gregariousness or extroversion might be considered important in the sense that extremely withdrawn or bashful children would be at a disadvantage in the classroom situation (Kohn, 1977). However, sociability outside the schoolroom -- specifically, participation in peer groups -- could plausibly have both positive and negative effects. For example, members are orginarily expected to conform to the group's norms to a high degree. If the group norms were anti-adult and anti-school, peer group participation would have negative implications for future school success. On the other hand, musical and athletic peer groups would involve systematic training and practice relevant to the participants' occupational goals (Schulz, 1966). It would seem useful to describe the extent and kind of social activities in which the children participated. However, lacking detailed prior evidence on the relationship of sociability to academic performance, it seems prudent to refrain from hypothesizing, a priori, treatment/control differences.

---

[1] Typically defined as the individual's striving to succeed, either in competition with others or in comparison to a set of standards (McClelland, 1955).

Methods

In order to assess the four areas of non-cognitive outcomes of preschool intervention programs, 15 items were selected from the Youth Interview and one item from the Parent Interview for use as dependent variables.[1] An additional six dependent variables were created, which were not independent of the original 16, for a total of 22. Table 16 provides a summary of the labels, sources, and coding of the dependent variables. Further coding details may be found in Appendix E.

These data were analyzed in two ways: (1) treatment (i.e. preschool vs. no preschool) main effects and (2) interactions of treatment with sex and with age on the dependent variables. The unit of analysis was individuals within projects. The association between treatment/control status and each of the dependent variables was calculated separately for each project.[1] Then, the results were pooled across projects by converting the p values to z-scores and summing the z-scores.

---

[1] For a more detailed description of the hypotheses, procedures and results, see Daniel M. Koretz, Long-term non-cognitive effects of seven infant and preschool intervention programs. Cornell University, University Microfilms, 1978.

The main treatment effects were scrutinized within sex and within age groups. For example, the various projects were dichotomized into an "old" group (subjects aged 15 to 19) and a "young" group (subjects aged 9 to 13) and the analysis performed separately to see whether the age groups differed. Interaction analyses of treatment by sex, treatment by age and treatment by sex by age were also performed within projects and the results then pooled.[2]

As is our common practice, the significance levels of statistical tests were adjusted to correct for large number of significance tests which were made. In this case, the correction factor was derived separately for each of the four domains:

---

[1] For dichotomous measures, the phi coefficient was used. For ordinal measures, the point-biserial correlation was used.

[2] For treatment by sex, a multiple regression was run with treatment, sex, and the product of treatment and sex as predictors. The significance of the beta weight of the product was the significance of the interaction. For treatment by age, the main effects of treatment were pooled across projects within each age group, yielding one pooled z score for the older projects and another for the younger projects. The interaction was then tested by the formula:

$$z = \frac{z_{old} - z_{young}}{\sqrt{2}}$$

Testing treatment by sex by age interactions involved combining the above two procedures. The within-sample sex by treatment interactions were pooled across each age group and the two resulting z scores used in the formula above. The last step in each of these procedures was simply to convert the z score to a p value.

achievement orientation (nine variables), family context of achievement orientation (four variables), self-evaluation (two variables) and sociability (seven variables).

The analyses were performed on the seven projects with the most complete data as of July, 1977: Beller, Gordon, Gray, Levenstein, Miller, Palmer, and Weikart.[1] A total of 702 Youth Interviews and 747 Parent Interviews were collected by these investigators.

Results

A broad picture of all children's aspirations (combining treatment and control groups) compared to their parents' positions in life provides a perspective on the treatment/control comparisons which follow. As Table 17 shows, the aspirations of these low-income, predominately Black children far exceeded their parent's attainments. Most children aspired to white collar jobs; their parents were largely semi-skilled or unskilled employees. The children planned (not hoped) to attend and to complete college; parents at most had graduated from high school. Interestingly, other investigators studying lower-class Black youngsters report similar high aspirations and expectations (cf. Spenner & Featherman, 1977). Similarly, all the children rated themselves somewhat better in school performance relative to their peers (see Table 20). In addition, 79% said they got along well with their families, 85% reported that they had "special friends" and 49% reported participating in organized community-wide activities.

---

[1] For various reasons (e.g., incomplete interview data from some projects), all seven were not necessarily included in each of the analyses.

Let us turn now to the consideration of treatment/control differences in the four areas of noncognitive outcomes.

Family Context of Achievement-orientation

A plausible explanation of enduring effects of preschool intervention programs is that parental expectations and behaviors were changed, enabling parents to further support and motivate their children. The four variables under "family context of achievement-orientation" provide a picture of one aspect of parental expectations: mothers' occupational aspirations for their children. Table 18 provides a summary of the results of comparing mothers of treatment (preschool program) children and the mothers of control children.

As Table 18A shows, there was a trend associating preschool attendance with higher maternal occupational aspirations for their children. In four of six projects treatment mothers had higher aspirations for their children compared to control mothers. In the two exceptional cases, the differences were essentially zero. The overall trend across the six projects was marginally significant (p = .065, two-tailed).

The strongest finding, however, involves variable 3; when asked "what kind of job would you like (your child) to have later in life," mothers of children who had attended preschool consistently responded with occupations which were higher than their children's answers to the same question.[1] Mothers of control children varied; some had higher occupational goals and some had lower goals for the child than the child had for himself. (For convenience, the difference between the mother's aspirations for the child and the child's own aspirations is called the "discrepancy score.") This treatment/control difference in discrepancy scores was statistically significant across projects -- $p = .0056$, two-tailed, or $p = .023$ after correction for multiple comparisons. It was also robust; deleting the strongest result (Gordon), the overall result was significant at the .031 level. These data are shown on Table 18B.

To summarize, there is some evidence that preschool changed the family context, and thus, perhaps, mother-child interactions, with respect to achievement orientation. Mothers of treatment children had higher vocational aspirations for the children than the children had for themselves. Additionally, mothers of treatment children tended to have higher aspirations for their children's future vocations than did mothers of control children.

Children's Achievement-Orientation

Before reporting the results of hypothesis-testing in the area of achievement orientation, the coding of "achievement-related" responses must be explained. The reader will note that variables

---

[1] See Table 19A for children's responses.

9, 10, 11, 12, and 13 on Table 16 are dichotomous variables with subjects' responses coded as either achievement-related or not achievement-related. Deriving a code proved difficult, primarily due to conceptual ambiguity of the construct of "achievement-orientation," which contains both attitudinal and motivational components (among others). It was not always possible to distinguish responses reflecting one aspect from those reflecting another. Therefore, responses were classified as "achievement-related" if they reflected either a high level of achievement motivation or a high value placed on presumably school-relevant activities and traits. The actual response categories and their coding may be found in Appendix E along with analyses of the construct validity of these particular variables.

The reader will recall the high educational expectations and occupational aspirations voiced by all the children. As Table 19 summarizes, there was no difference between treatment and control children's educational expectations. Treatment children in older projects showed a slight tendency (p = .091) to report lower vocational aspirations (Table 19A). There were no differences in the areas of employment and leisure activities.

The strongest finding was for attitudes toward the self; when asked to "tell me something you've done that made you feel proud of yourself", children with preschool experience were far more likely to respond with achievement-related reasons (such as school or job achievements, straightening oneself out, helping out at home[1]) rather than other reasons (such as good behavior or altruistic acts). This result was significant -- p = .0025, two-tailed or p = .023 when corrected for multiple comparisons. It was also robust (i.e., still significant when the project with strongest results was removed). It is also consistent across projects (p = .032, two-tailed, by the sign test). Table 19B presents these data. As is evident, the younger projects and older projects differed only slightly. However, breaking down the sample further by sex as well as by age revealed an interesting difference. As Table 19C shows, the treatment effect was similar for males and females in the younger projects. For older projects the treatment effect was significant for females (p = .005) but not for males (p = .254). Although our complex pooled-z significance tests were not performed on individual items, treatment girls were more likely to name school-related reasons or helping at home as bases for feeling proud of themselves. In contrast, control girls more frequently said they were proud of having babies or said they had no reason to be proud of themselves.

In summary, some areas of achievement motivation appear to be affected by preschool experience and others do not. The strongest treatment effect was on children's attitudes toward themselves; children who participated in preschool were more likely to report feeling proud of themselves for achievement-related reasons than were control children. This was especially true for girls. The result is particularly interesting since the question, "what makes you feel proud of yourself?" represents something of an interface between achievement motivation and self-concept.

---

[1] Helping out at home was coded as achievement-related because it ordinarily constitutes an onerous chore and does represent work activity as opposed to non-work responses such as good behavior, etc. For relevant construct validity analyses, see Appendix E.

Children's Self-evaluations

Children normally engage in the process of comparing themselves to others and judging their own prowess and standing relative to their peers. There were two questions on the Youth Interview which attempted to tap such social comparisons as a means of measuring self-esteem. One asked the children to judge their own academic performance relative to others in their class and the other asked how well they got along with the other members of their household.

Table 20 presents the data concerning children's evaluations of their own academic performance. It is noteworthy that both treatment and control children rated themselves somewhat above their peers. While there was no overall treatment-control difference, there was a small difference in the older projects: treatment subjects rated themselves higher than did controls. Although the difference was not significant in any one project alone, it was significant when the three older projects were taken together (p = .039).[1]

Analysis of the children's self-evaluation of getting along with other household members revealed no treatment-control differencces.

In sum, only two interview items were categorized as self-evaluation items. One revealed no treatment effects on children's self-appraisals of their compatability with their families while the other indicated that treatment children in the older projects had slightly but significantly more favorable self-evaluations of their own academic performances. These results are summarized in Table 21.

Children's Sociability and Social Participation

The dependent variables included in these analyses are presumed to tap different aspects of this general domain. Variables 16 and 17 (the amount of participation in organized community activities) and variables 18 and 19 (the amount of contact with "special friends") were used to describe patterns of social interactions. As the summaries on Table 21 show, none of the treatment-control differences was striking in this domain. In the area of patterns of social interaction, approximately 85% of both treatment and control children reported participating in organized community activities. There was no significant difference between treatment and control children in reporting the existence of special friends (Table 21A). Among children in the older projects, there was a marginally significant trend (p = .069) for treatment children to report spending less time with special friends compared to controls (Table 21B).[1] Taken together, the results of these four variables suggest that while treatment children report themselves to be as socially active in organized community activities as control children, they (especially girls)[2] may tend to spend less time with close friends. Some might interpret this positively, as an indication of social autonomy. However, Ladner (1971) argues persuasively that female adolescent peer groups provide vital sources of emotional support and positive socialization in low-income Black communities.[1] As mentioned earlier, there is no detailed prior evidence on the relationship of sociability to academic performance, further complicating interpretative efforts. Thus, we report this slight trend as part of our exploratory research but will refrain from theoretical interpretations.

---

[1] The small size of differences and the problem of multiple comparisons make it seem likely that the difference is merely due to chance. However, further scrutiny of the data suggests that there is a true age difference. In all the younger projects, controls rate themselves more highly than do treatment children and the reverse is true in all older projects. Accordingly, the treatment by age interaction is significant at the .03 level, two-tailed. Such a clear association between age and the direction of the differences is unlikely to occur by chance alone (p = .057, two-tailed, by Fisher's exact 2 x 2 test). Essentially identical age patterns appeared in the construct validity analyses (Appendix E), lending further credence to the view that this difference is developmental rather than due to chance.

---

[1] If broken down further by sex, this treatment-control difference was marginally significant among females (p = .052, two-tailed). The sex by treatment interaction, however, was not significant.

[2] If broken down further by sex, these data indicate that treatment females tended to be less likely to report having special friends (p = .072, two-tailed). Treatment-control differences were small and inconsistent among males. The sex by treatment interaction however, was not significant.

Variables 20 and 22 (sociability ratings of school and spare-time activities) were used as measures of social attitudes and, indirectly, of children's own gregariousness.

Treatment-control differences were again small or nonexistent. When asked to tell "what is the best thing about school," children who had attended preschool were more likely than controls to respond with sociability-related answers, such as being with friends and meeting new people (p = .041).[2] Table 21C contains these data. Table 21D shows a non-significant trend (p = .081, two-tailed) for treatment children to respond to "what do you do in your spare time" with answers such as "play with friends" or "talk on the phone" compared to children in control groups. These results are only trends; but it seems fair to conclude on the basis of these variables that preschool experience apparently did not impair children's sociability.

---

[1] Stack (1974) also documents the crucial importance of female ties with kin and fictive kin ("special friends") in the social organization of lower-class Black communities.

[2] This result should be interpreted cautiously. Sociability-related responses were relatively infrequent across all projects; thus, some of the scores reflect the responses of only a few individuals. Furthermore, the strongest results were from the Levenstein project, which had a self-selected control group. This, in conjunction with the failure to remain significant after correction for multiple comparisons, suggests that the result should be seen as a suggestive trend only.

Summary and Conclusions

These analyses of non-cognitive outcomes were exploratory: we knew preschool had positively affected children's school performance and that it affected more than just cognitive ability. The question was: what was affected? Investigators in the 1960s suggested it would be important to influence children's achievement motivation and self-esteem and to affect parents' attitudes as well. As detailed above, there is some indication that preschool affected those areas. For treatment but not control children, mothers' aspirations for the children were higher than children's aspirations for themselves. This and other trends suggest that preschool may have affected the familial context with respect to achievement orientation. Compared to control children, preschool children were more likely to give achievement-related reasons for being proud of themselves. This was especially true for girls. Older treatment children rated themselves as superior to their classmates in school performance and tended to have lower (and, in this context, more realistic) vocational aspirations.

Although many of the other results were only trends, the reader should remember that these measures were administered 10 to 15 years after the children participated in preschools. In view of this, these findings lend credence to the hopes that attitudes could be changed and that such changes would persist.

Head Start and other programs were initiated in the hope that changing children's abilities and attitudes would be instrumental in extracting them from poverty. This appears naive in retrospect. Unemployment for Black youngsters (approximately 90% of our subjects) was 39% in 1976 and has not substantially improved to date. Furthermore, Black unemployment for all ages has been twice that for whites since World War II (Kenniston, et al., 1977). The relationship between level of educational attainment and later vocational status has historically been very low for Black persons. For example, a college-educated Black male can expect to attain the average job and income level of a white male with no more than a high school diploma (Ogbu, 1978). It seems important to recognize that preschools can make a difference, but that the larger society and its institutions must also change in order to fulfill the promise of preschool for low-income children and their families.

Table 16

Dependent Variables for Analyses of Non-cognitive Outcomes

Variable	Question # on Interview[a]	Code[b]
A. Family context of child's aspirations		
1. Mother's occupational aspirations for own child	#16(P)	Seven point Hollingshead Scale
2. Percentage of mothers stating that child's vocation is up to him/her to decide	#16(P)	Percent of all mothers in group
3. Discrepancy between mother's and child's vocational aspirations	#16(P) #16(Y)	Algebraic discrepancy score
4. Absolute value of the discrepancy between mother's and child's vocational aspirations	#16(P) #16(Y)	Absolute value
B. Achievement-orientation		
5. Educational expectations	#2(Y)	1 = complete grammar school to 8 = graduate or professional training
6. Occupational aspirations	#16(Y)	Seven point Hollingshead Scale
7. Whether and where child is paid employee	#15a(Y)	1 = no paid work; 2 = paid work at home; 3 = paid work outside home
8. Amount of paid work	#15c(Y)	1 = no paid work to 4 = full time work
9. Achievement-orientation of spare-time activities	#8(Y)	1 = achievement-related; 0 = other (multiple responses summed)
10. Achievement-orientation rating of "best things about school"	#3(Y)	1 = achievement-related; 0 = other (multiple responses summed)
11. Achievement-orientation ratings of "worst things about school" (worst things = achievement)	#4(Y)	-1 = achievement-related; 0 = other (multiple responses summed)

Table 16 (Cont.)

Variable		Question # on Interview[a]	Code[b]
	12. Achievement-orientation ratings of reasons for being proud of oneself	#1[illegible](Y)	-1 = achievement-related; 0 = other
	13. Achievement-orientation ratings of reasons for admiring the most admired adult	#23(Y)	-1 = achievement-related; 0 = other
C.	Self-evaluation		
	14. Self-evaluation of academic performance	#5(Y)	1 = much worse than others to 5 = much better than others
	15. Self-evaluation of household harmony	#22(Y)	1 = get along with them badly to 5 = get along very well
D.	Sociability and Social Participation		
	16. Participation in organized community activities	#6(Y)	1 = yes; 0 = no
	17. Frequency of participation in organized activities	#6(Y)	Number of times per month
	18. Has "special" friend(s)	#7(Y)	1 = yes; 0 = no
	19. Frequency of contact with "special" friend(s)	#7(Y)	Number of times per month
	20. Sociability ratings of "best things about school"	#3(Y)	1 = sociability-related; 0 = other (multiple responses summed)
	21. Sociability ratings of "worst things about school"	#4(Y)	-1 = sociability-related; 0 = other (multiple responses summed)
	22. Sociability ratings of spare time activities	#8(Y)	1 = sociability-related; 0 = other

---

[a] (Y) = Youth Interview; (P) = Parent Interview. See Appendix for Youth and Parent Interviews.

[b] See Appendix E for details on coding decisions for these items.

Table 17

Educational and Occupational Attainment of Parents and Aspirations of Children, Treatment and Control Pooled*

	Mean	Mode
Father's Education	10.1 years	12 years
Mother's Education	10.4 years	12 years
Child's Educational Aspirations	some college	complete college
Father's Occupation (Hollingshead)	6.0	7
Mother's Occupation (Hollingshead)	6.7	7
Child's Desired Occupation (Hollingshead)	3.2	2
Mother's Desired Occupation for Child (Hollingshead)	2.7	2

Key to Hollingshead Occupational Codes:

1. Higher executives, proprietors of large concerns, and major professionals.
2. Business managers, proprietors of medium-sized businesses, and lesser professionals.
3. Administrative personnel, small independent business, and minor professionals.
4. Clerical and sales workers, technicians, and owners of "little businesses."
5. Skilled manual employees.
6. Machine operators and semi-skilled employees.
7. Unskilled employees.

(Hollingshead, 1957)

---

*Data on parents are from the time of the child's entry into the intervention program.

Table 18

Summary of Treatment Effects on Family Context of Achievement Orientation

Variable	Difference (2-tailed significance level uncorrected for multiple comparisons)
1. Mother's occupational aspirations for their children	Trend for treatment mothers to have higher aspirations (p = .065)
2. Percent of mothers stating that their child's vocation is the child's own decision	No significant difference by parametric methods; however, difference favored the treatment mothers in 5 of 6 projects (p = .094, by Wilcoxon signed-ranks test).
3. Discrepancy between mother's aspirations for child and the child's own aspirations	Treatment mothers' aspirations are higher, relative to those of their children than are those of control mothers relative to their children's (p = .0056). (.023 when corrected for multiple comparisons.)
4. Absolute discrepancy scores	No difference.

Table 18A

Mean of Mother's Occupational Aspirations for Child, Treatment vs. Control (Hollingshead Code)*

Project (n)	All Children	Treatment	Control	Difference	r_{pb}	p (2-tailed)
Young Projects (ages 9-13)						
Gordon (68)	2.57	2.54	2.67	-.13**	-.038	.756
Levenstein (64)	2.42	2.24	3.00	-.76	-.225	.074
Miller (132)	2.73	2.74	2.70	.04	.009	(.914)#
Palmer ‡ (58)	2.02	1.83	2.75	-.94	-.245	.064
Mean	2.44	2.34	2.78	-.44		.055
Old Projects (ages 15-19)						
Beller (72)	2.83	2.88	2.80	.08	.025	(.836)
Gray (34)	3.59	3.38	3.92	-.54	-.158	.372
Weikart ‡‡	--	--	--	--	--	--
Mean	3.21	3.13	3.36	-.23		.626
All Projects	2.69	2.60	2.97	-.37		.065

* Lower numbers correspond to higher aspirations; see Table 17 for key.

** A negative difference score indicates that the treatment mothers' aspirations are higher.

‡ All male sample.

‡‡ Data not usable.

# Figures in parentheses are in the reverse direction.

Table 18B

Mean Discrepancy Between Mother's Occupational Aspirations for Child and Child's Own Aspirations, Treatment vs. Control (Based on Hollingshead Codes)*

Project (n)	All Children	Treatment	Control	Difference	r_{pb}	p (2-tailed)
Young Projects (ages 9-13)						
Gordon (57)	-.70**	-1.02	.20	-1.22***	-.260	.050
Levenstein (41)	-.41	- .52	-.21	- .31	-.103	.520
Miller (125)	-.64	- .60	-.76	.16	.033	(.712)#
Palmer ‡ (53)	-.58	- .80	.17	- .97	-.241	.082
Mean	-.58	- .74	-.15	- .59		.048
Old Projects (ages 15-19)						
Beller (64)	-.31	- .77	-.07	- .70	-.186	.142
Gray (32)	-.31	- .70	.33	-1.03	-.247	.172
Weikart ‡‡	--	--	--	--	--	--
Mean	-.31	- .74	.13	- .87		.047
All Projects	-.49	- .74	-.06	- .68		.0056

* See Table 17 for Hollingshead Code.
These results are based on a smaller sample than Tables 18A and 19A, since all cases in which either the child's aspirations or the mother's aspirations were missing had to be dropped for this analysis.

** A negative score indicates that the children aspired to a lower level of status than did their mothers.

*** A negative difference score indicates that the aspirations of the treatment children are lower, relative to their mothers', than the control children's are, relative to their mothers'.

‡ All-male sample.

‡‡ Data not usable.

# Figures in parentheses are in the reverse direction.

Table 19

Summary of Treatment Effects on Achievement-Orientation Variables

Variable	Difference (2-tailed significance level uncorrected for multiple comparisons
5. Children's educational expectations	No difference
6. Children's occupational aspirations	No significant difference across all projects. Trend in older projects for treatment children to have lower aspirations (p = .091).
7. Whether and where employed	No difference
8. Amount of paid work	No difference
9. Achievement-orientation ratings of spare-time activities	Control greater than treatment (p = .044). No difference if athletics not classified as achievement-related.
10. Achievement-orientation ratings of "the best thing about school"	No significant differences across all projects. Trend in younger projects for treatment children to have higher ratings than control, (p = .063).
11. Achievement-orientation ratings of "the worst thing about school"	Parallels #10, but not significant.
12. Achievement-orientation ratings of reasons for being proud of oneself	Treatment greater than control (p = .0025). (.023 when corrected for multiple comparisons.)
13. Reasons for admiring most admired adult	No difference

Table 19A

Mean of Children's Occupational Aspirations, Treatment vs. Control (Hollingshead Code)[*]

Project (*n*)	All Children	Treatment	Control	Difference	r_{pb}	*p* (2-tailed)
Young Projects (ages 9-13)						
Gordon (81)	3.26	3.42	2.65	.77[**]	.176	.116
Levenstein (72)	3.15	3.10	3.25	-.15	-.043	(.720)[#]
Miller (134)	3.38	3.35	3.48	-.13	-.032	(.710)
Palmer [‡] (130)	2.83	2.78	3.00	-.22	-.055	(.538)
Mean	3.16	3.16	3.10	.06		.91
Old Projects (ages 15-19)						
Beller (93)	3.11	3.52	2.88	.64	.178	.088
Gray (49)	3.80	3.94	3.53	.41	.108	.462
Weikart (63)	3.75	3.89	3.64	.25	.065	.616
Mean	3.55	3.78	3.35	.43		.091
All Projects	3.33	3.43	3.20	.23		.231

[*] Lower numbers correspond to higher aspirations; see Table 17 for key.

[‡] All male sample.

[**] A positive difference score indicates that the treatment children's aspirations are lower.

[#] Figures in parentheses are in reverse direction.

Table 19B

Percent of Children Giving Achievement-Related Reasons for Being Proud of Themselves, Treatment vs. Control

Project (n)	All Children	Treatment	Control	Difference	Φ	p (2-tailed)
Young Projects (ages 9-13)						
Gordon (85)	85.9%	88.2%	76.5%	11.7%	.135	.218
Levenstein ‡‡	--	--	--	--	--	--
Miller (141)	77.3	78.9	71.9	7.0	.070	.408
Palmer ‡ (102)	72.5	77.8	52.4	25.4	.230	.020
Mean	78.6	81.6	66.9	14.7		.012
Old Projects (ages 15-19)						
Beller (95)	62.1%	65.7%	60.0%	5.7%	.057	.584
Gray (50)	68.0	75.8	52.9	22.9	.232	.106
Weikart (64)	81.3	86.2	77.1	9.1	.116	.364
Mean	70.5	75.9	63.3	12.6		.078
All Projects	74.5%	78.8%	65.1%	13.7%		.0025

‡ All male sample.

‡‡ Data not available.

Table 19C

Percent of Children Giving Achievement-Related Reasons for Being Proud of Themselves, by Sex, Age, and Treatment/Control Status

		Treatment	Control	Difference
Young Projects ‡‡ (ages 9-13)	Males (including Palmer ‡)	82.9%	72.2 %	10.7 %
	Males (excluding Palmer)	85.5	82.15	3.35
	Females	81.9	68.35	13.55
Old Projects (ages 15-19)	Males	74.7%	75.4 %	-00.7 %
	Females	77.6	53.5	24.1

‡ Palmer's sample is all male.

‡‡ Data not available for Levenstein's project.

Table 20

Mean of Children's Self-Evaluations of Their Overall School Performance Relative to Their Peers, Treatment vs. Control*

Project (n)	All Children	Treatment	Control	Difference	r_{pb}	p (2-tailed)
Young Projects (ages 9-13)						
Gordon (92)	2.78	2.81	2.70	.11	.053	(.614)#
Levenstein (75)	2.63	2.66	2.56	.10	.051	(.662)
Miller (141)	2.84	2.86	2.78	.08	.045	(.596)
Palmer ‡ (144)	2.67	2.69	2.59	.10	.052	(.538)
Mean	2.73	2.76	2.66	.10		(.301)
Old Projects (ages 15-19)						
Beller (96)	2.63	2.43	2.74	-.31**	-.167	.104
Gray (50)	2.74	2.70	2.82	-.12	-.115	.428
Weikart (69)	2.64	2.52	2.74	-.22	-.144	.236
Mean	2.67	2.55	2.77	-.22		.039
All Projects	2.70	2.67	2.70	-.03		.566

* 1 = Much better than others in your classes; 3 = About the same as others; 5 = Much worse than others.

‡ All male sample.

* A negative difference score indicates that the treatment group's self-evaluations were higher than those of the controls.

# Figures in parentheses reflect negative findings.

Table 21

Summary of Treatment Effects on Self-Evaluation and Sociability Variables

Variable	Difference (2-tailed significance level, uncorrected for multiple comparisons)
Self-Evaluation Variables	
14. Self-evaluation of academic performance	No significant difference across all projects. Treatment higher than control in older projects (p = .039)
15. Self-evaluation of how well one gets along with those with whom one lives	No difference
Sociability Variables	
16. Participation in organized community activities	No difference
17. Frequency of participation in organized community activities	No difference
18. Whether one has "special" friends	Non-significant trend (p = .12) for treatment children to be less likely to report having special friends. Marginally significant among females (p = .072); inconsistent and non-significant among males
19. Frequency with which one gets together with special friends	No difference across all projects; marginally significant (p = .069) trend toward lower frequencies among treatment children in the older projects. Marginally significant among (all) females (p = .057), but inconsistent and non-significant among males
20. Sociability ratings of the "best things about school"	Treatment greater than control (p = .041) {.246 when corrected for multiple comparisons}
21. Sociability ratings of the "worst things about school"	No variance
22. Sociability ratings of spare time activities	Treatment marginally greater than control (p = .082)

Table 21A

Percent of Respondents Who Stated that They Have "Special Friends" With Whom they Spend Time, Treatment vs. Control

Project (n)	All Children	Treatment	Control	Difference	Φ	p (2-tailed)
Young Projects (ages 9-13)						
Gordon (92)	91%	90%	95%	- 5.0%	-.069	.512
Levenstein (74)	93	92	96	- 4.0	-.072	.544
Miller (141)	91	89	97	- 8.0	-.114	.178
Palmer ‡ (144)	87	88	81	7.0	.088	(.296)#
Mean	90.5	89.8	92.3	- 2.5		.439
Old Projects (ages 15-19)						
Beller (96)	82%	71%	89%	-18.0%	-.216	.034
Gray (50)	78	82	71	11.0	.128	(.374)
Weikart (72)	76	69	82	-13.0	-.161	.178
Mean	78.7	74.0	80.7	- 6.7		.137
All Projects	85.4%	83.0%	87.3%	- 4.3%		.119

‡ All male sample.

# Figures in parentheses are in the reverse direction.

Table 21B

Mean Frequency with which Respondents Spend Time with "Special Friends" (Times per month), Treatment vs. Control

Project (n)	All Children	Treatment	Control	Difference	r_{pb}	p (2-tailed)
Young Projects (ages 9-13)						
Gordon (92)	16.90	16.78	17.37	- .59	-.047	.656
Levenstein (75)	17.39	17.28	17.60	- .32	-.025	.834
Miller (140)	17.26	17.11	17.75	- .64	-.042	.622
Palmer ‡ (112)	14.95	15.39	13.33	2.06	.103	(.278)#
Mean	16.63	16.64	16.51	.13		.981
Old Projects (ages 15-19)						
Beller (96)	14.88	12.80	16.07	-3.27	-.195	.056
Gray (50)	11.12	10.88	11.59	- .71	-.040	.782
Weikart (70)	13.20	12.06	14.16	-2.10	-.118	.330
Mean	13.07	11.91	13.94	-2.03		.069
All Projects	15.10	14.61	15.41	- .80		.227

‡ All male sample.

# Figures in parentheses are in the reverse direction.

Table 21C

Mean Sociability Ratings of "The Best Things about School,"[*] Treatment vs. Control

Project (*n*)	All Children	Treatment	Control	Difference	r_{pb}	*p* (2-tailed)
Young Projects (ages 9-13)						
Gordon (88)	.15	.14	.16	-.02	-.015	(.890)[#]
Levenstein (75)	.20	.30	.00	.30	.289	.012
Miller (141)	.05	.06	.00	.06	.124	.144
Palmer ‡ (144)	.10	.10	.13	-.03	-.034	(.686)
Mean	.13	.15	.07	.08		.086
Old Projects (ages 15-19)						
Beller (96)	.26	.34	.21	.13	.142	.166
Gray (50)	.00	.00	.00	.00	--	--
Weikart (66)	.03	.04	.03	.01	.027	.830
Mean (excluding Gray)	.15	.19	.12	.07		.263
All Projects (excluding Gray)	.13	.16	.09	.07		.041

* 1 = sociability-related; 0 = other (multiple responses summed).

‡ All male sample.

# Figures in parentheses are in the reverse direction.

Table 21D

Mean Sociability Ratings of Spare-Time Activities, Treatment vs. Control*

Project (n)	All Children	Treatment	Control	Difference	r_{pb}	p (2-tailed)
Young Projects (ages 9-13)						
Gordon (87)	.48	.46	.58	-.12	-.090	(.406)[#]
Levenstein ⌗	--	--	--	--	--	--
Miller (141)	.38	.38	.38	.00	.0008	.922
Palmer ‡ (144)	.33	.37	.19	.18	.154	.066
Mean	.40	.40	.38	.02		.562
Old Projects (ages 15-19)						
Beller (95)	.51	.47	.52	-.05	-.039	(.708)
Gray (48)	.08	.13	.00	.13	.223	.128
Weikart (96)	1.08	1.23	.96	.27	.216	.034
Mean	.56	.61	.49	.12		.061
All Projects	.48	.51	.44	.07		.081

* 1 = sociability-related; 0 = other (multiple responses summed).

‡ All male sample.

⌗ Data not available.

[#] Figures in parentheses are in the reverse direction.

Use of Title IV Child Welfare Services

As pointed out earlier, Head Start included in its goals the improvement of family functioning. Most of the preschool programs comprising the Consortium predated Head Start, but they, too, were concerned to some degree with family welfare. As a result, we undertook an exploratory study to determine whether there would be treatment/control differences in the families' use of Title IV child welfare services.[1] Four of the Consortium projects agreed to take part in this follow-up. Because the agencies providing service data were assured that the reporting of this research would protect specific information sources, the projects will not be identified by name.

Methods

The gathering of data for this study served to test the feasibility of using archival records to measure program impacts. Archival records are less costly to obtain than data generated by observation, testing or interviewing. Each of the four participating projects was asked to submit to the Cornell staff one list of treatment children's names, addresses, birth dates, and parents' names and a separate list of control children's names, etc. Any identifying labels were removed. In each of the four corresponding states, the director of the division responsible for the administration of Title IV child welfare services was then contacted and sent a request form.

---

[1] For more details, see Marilyn Rosche, Early intervention and later use of child welfare services. Unpublished Master's Thesis, Cornell University, 1979.

The form asked for the total number of children from each list who had received any child welfare service (overall use), and then asked that the figures be broken down for each type of service provided. Thus, any child who had received multiple services would be counted several times under the breakdown by service type but only once in the overall usage tally.[1] Thus, the Cornell staff received from each agency data pertaining to two lists -- one for treatment and one for control children -- making a treatment/control comparison possible. This procedure circumvented the confidentiality issue. Since none of the service information gathered could be linked to a particular subject, it was not necessary to obtain the subjects' consent to have the information released. Considering the financial and time limitations ordinarily imposed on research efforts, such an approach enables investigations which otherwise might be impossible to consider.[2]

The available data differed somewhat for each of the four projects as follows. The preliminary contact with State A revealed that the requested information was not available at the state level but could be obtained through the appropriate county offices. Since all of the subjects were originally located in one county, this presented no problem. State agency personnel forwarded the request to county staff, who sent their responses directly to Cornell. However, any subjects who might have moved to another county and subsequently received services would not be included.[1]

---

[1] Information was requested for services provided over the past 5 years rather than at any time during the child's life on the assumption that such a request would be more favorably received than one requiring a more extensive search.

[2] The data do not provide as much information, of course, as would a first-hand scrutiny of individual case records. For instance, this research does not permit us to determine whether those children who received child welfare services are the same individuals who failed in school or were placed in special education classes. However, an indication of treatment/control differences in service usage might encourage individual project staff to obtain subjects' consent to examine case records.

Personnel from State B arranged to send the available information to Cornell through their computerized case-tracking system. Based on individual client transactions as reported by caseworkers, this information is used not only by the division providing child welfare services but by the courts and residential facilities as well. Rather than including all services which might be categorized as child welfare, the data received by Cornell were limited to records of substitute care arrangements and accompanying counseling. Since the computerized system began in 1976, the accuracy of the data is questionable prior to that time. These data, then, must be considered an underestimate of service usage over the past 5 years. In addition, it was difficult to differentiate between various types of substitute care arrangements based on the information provided.

State C also used a computerized data system to gather the requested information. Their response to Cornell included usage figures for a dozen service categories as well as overall use. The data were generated from computer file records which were activated in 1975. Cases initiated prior to the change over to this system were included providing the case was listed as "active" in 1975. Cases initiated and terminated before that time would not have appeared in the read out.

---

[1] Although this could conceivably result in a lower proportion of subjects reported as having received services, there is no reason to believe that it would differentially affect the experimental and control groups.

Since child welfare services in State D are locally administered, contact was made with the appropriate city agency rather than the state. However, the city agency was unable to provide the requested information. As an alternative, contact was made with a private information system which gathers data on foster care placements. This agency was able to provide data on current foster care placements for the children or their siblings.

Results

Project A. The response from State A included data for total service use, plus a breakdown into 11 service categories: adoption, day care, family and individual counseling, family planning, health related services, housing improvement, information and referral, interstate/intercounty, protective services for children, services for unmarried parents, and transportation services. Comparisons of the frequency of service use revealed no significant differences between experimentals and controls. The proportion of the sample receiving any service -- approximately 50% for both treatment children and controls -- was far higher than for the other projects.

Project B. In addition to the figures reflecting overall service use, State B provided data for five types of substitute care arrangement plus counseling services. None of the experimental children were reported as having received any services; six control children were included in the overall service category. A chi-square test of significance showed this difference to be significant at the .05 level (two-tailed). The treatment-control difference in use of foster care services was also significant ($p < .05$, two-tailed). None of the comparisons in the other service categories (adoption, counseling, halfway house, institutions, shelter care) reached significance.

Project C. The response from State C included figures for overall service use plus a breakdown into 12 service categories: child protection services, counseling for emotionally disturbed youth, education and training services, family counseling services, family planning, foster care, health related services, home management, housing and home improvement, juvenile services in the community, social service planning services, and unmarried parent services. Although proportionately more experimentals than controls were reported as having received any type of service (overall use), this difference did not approach significance. None of the treatment/control comparisons in the 12 service categories were significant.

Project D. Two categories of service status were reported by the data system contacted for information: the first indicated that the subject was under foster care; the second indicated that the subject was not under care himself but one or more of his siblings were. No other service categories were included. None of the control children appeared in either category; four experimental children were under care and an additional two experimental children had siblings under care. These differences were not statistically significant.

In general, participation in preschool intervention programs apparently did not affect the incidence of use of Title IV child welfare services. However, the data collection method devised for the study proved workable. It avoided problems of confidentiality and was both relatively rapid and inexpensive. Such an approach may prove fruitful in future explorations of program outcomes.

DIFFERENTIAL EFFECTS OF PRESCHOOL

Which Children Benefit Most on WISC-R Test Scores?

The reader will recall from results reported earlier in this report that in comparisons of treatment and control children on the WISC-R IQ tests (administered in 1976-77) only Levenstein and Palmer found differences on the full, verbal, and/or performance scores and the subtests. Further analyses seemed warranted, however, to check the possibility that interaction effects were present; that is, that certain kinds of children (e.g., boys vs. girls) had benefitted from preschool while others had not. On the basis of earlier work (Murray, 1977) we chose to investigate sex of child and mother's level of education. Mother's education had a strong positive correlation with the children's WISC-R scores and with earlier Stanford-Binet IQ scores. Sex was linked to different patterns of relationship between IQ and various demographic variables (Hubbell, 1977; Murray, 1977).

Work by other investigators has found differential effects of these variables. For example, Black females scored much higher than males on IQ and scholastic achievement tests. As adults, Black women had more education, were more likely to fall in the highest occupational categories and had higher income compared to Black males (Jensen, 1971; Sowell, 1973).

The relationship between socioeconomic status and various measures of academic ability and performance has been extensively documented (cf. Boocock, 1972). Family SES is usually measured by three status variables: (a) level of parental occupation(s); (b) level of education attained by parents; and (c) level of family income. Most professionals would agree that these and other measures of SES are merely convenient indicators of the sociopsychological environment of the family -- aspects such as power relations, values, interactional patterns, parental expectations, etc. These psychological correlates of SES probably link parental SES to children's achievement. Lacking more refined measures of the family environment, we must depend on the indicator variables for all analyses in this report. Since almost half of the families in our data bank were single-parent families and since many families derived their income from welfare payments (with payment levels varying across states), we decided to use mother's educational level as the best indicator of the family SES. In addition, mother's education is of interest because there could be a direct relationship between the mother's education and the children's IQ and academic performance. Since mothers are generally the primary parental caretakers of children under age 5, their influence may be correspondingly strong. And it is at least feasible that more educated mothers impart different values, attitudes and motivations relevant to school success to their children than do less educated mothers.

Methods

Once again, as was true for the earlier analysis of WISC-R, the projects were analyzed individually and results were not later pooled. First the effects of sex were examined by investigating the correlation between sex and the WISC full, verbal, performance and subtest scores. Then an analysis of variance was performed to investigate the combined effect of preschool attendance and sex.[1]

Mother's level of education (as reported before children began preschool programs) was first correlated with the WISC-R scores.[2] The combined effect of preschool attendance and mother's education was investigated by using multiple regression. The linear terms of preschool vs. no preschool and level of mother's education were entered and then their crossproduct was entered into the equation.[3]

Results

First we will consider the two interaction analyses. An ANOVA was performed with the factors of sex and preschool attendance. There were no significant interactions on either the full, verbal or performance IQs or the subtest scores. Thus, there was no indication that early childhood education had differential effects with respect to sex on WISC-R full or subtest scores.

---

[1] Palmer had only male subjects and thus was not included in this analysis.

[2] For the WISC analysis only, the variable of mother's education is different for Gordon's project. For WISC analysis, the mother's educational level reported in the 1976 Parent Interview was used. Since the Gordon project specifically encouraged mothers to go back to school, it is no surprise that mothers of program children reported significantly higher mean education after as compared to before the program (pre-program mean = 10.69; 1976 mean = mean = 11.54 years; t = 3.69, p = .001). In interpreting Gordon's data, one should remember that the effects of preschool education and mother's education may be confounded for WISC analysis.

[3] Had any of the above analyses warranted, all three variables -- preschool attendance, sex, and mother's education -- would have been included in a multiple regression equation.

A multiple regression analysis was performed using the linear terms of maternal education and preschool attendance and then entering their crossproduct in the equation. There were no significant interaction effects for mother's education and preschool attendance on WISC-R full or subtest scores. Thus, there was no evidence that preschool differentially affected the IQ scores of children whose mothers had completed different years of education.

In the process of performing the interaction analyses, both sex and maternal education were correlated with the WISC-R scores. The results of those correlations are presented briefly here for interested readers.

Regarding the correlation between sex and WISC-R scores, only the Weikart project showed a significant correlation. As Table 22 shows, boys scored higher than girls on full IQ, verbal IQ and performance IQ. This relationship disappeared, however, upon correcting for multiple comparisons. On subtest scores (Table 23), boys scored significantly higher than girls in three of the five projects on Picture Completion, with boys in Miller's project showing a similar tendency. The most consistent significant correlations were found in the coding B subtest; girls performed better than boys in all projects. Other subtests revealed no significant correlations after correction for multiple comparisons.

As Table 24 shows, there was a positive correlation between mother's education and verbal IQ in all six projects. Three projects -- Gordon, Palmer, and Weikart -- found a positive correlation between mother's education and performance IQ. Considering the subtests (Table 25), Vocabulary had the strongest relationship to mother's education (five projects significant; three significant after adjustment for multiple comparisons) with less consistent patterns of significant positive correlation for the other verbal subtests. The only performance subtests with reasonably consistent positive correlations were Digit Span and Coding.

In summary, only a few subtests correlated significantly with sex. In contrast, there was a pattern of significant positive correlations between mother's education and WISC-R scores, especially on verbal subtests. Despite these correlations, there were no significant interactions. When WISC-R scores were used as the outcome measure, there was no evidence that preschool benefitted girls more than boys or vice-versa or that children whose mothers had different levels of education benefitted more or less from attending preschool.

Table 22

Correlation of Sex with WISC-R IQ Scores, by Project

Project (n)		Correlation	Significance (two-tailed)
Gordon (90)	FIQ[a]	.0890[d]	.202
	VIQ[b]	.1709	.054
	PIQ[c]	-.0240	.411
Gray (52)	FIQ	-.0369	.398
	VIQ	-.0772	.293
	PIQ	.1198	.199
Levenstein (76)	FIQ	-.0212	.428
	VIQ	-.0691	.277
	PIQ	.0372	.375
Miller (141)	FIQ	.0358	.337
	VIQ	.0766	.183
	PIQ	-.0167	.422
Weikart (110)	FIQ	.2100	.014*
	VIQ	.2044	.016*
	PIQ	.1674	.040*

* $p < .05$ before correction for multiple comparisons.

[a] Full IQ score.

[b] Verbal IQ score.

[c] Performance IQ score.

[d] Positive correlations indicate boys scored higher than girls. Negative correlations indicate girls scores higher than boys.

Table 23

Correlation of Sex with WISC Subtest Scores, by Project

Project	Subtest[a] 1	2	3	4	5	6	7	8	9	10	11	12	13
Gordon													
Correlation[b]	.1147	.1132	.1137	.0740	.0746	--	.1208	-.0809	.1372	-.0300	--	-.2805	--
Significance	.145	.148	.147	.248	.246	--	.133	.228	.103	.391	--	.004*	--
Gray													
Correlation	.0888	-.0192	.0148	-.0887	-.2783	--	.3025*	.2040	.0609	.2105	-.2091	--	--
Significance	.266	.446	.459	.266	.023*	--	.015	.073	.334	.067	.068	--	--
Levenstein													
Correlation	-.0600	.0258	-.2558	.0948	.0737	-.1245	.2385	.0424	.1627	.1107	--	-.3194	--
Significance	.303	.412	.013*	.208	.264	.142	.019*	.358	.080	.171	--	.002*	
Miller													
Correlation	.0358	-.0599	.0705	.0293	.0253	-.0156	.1357	-.0008	.1205	.1176	--	-.4081	.1008
Significance	.047*	.240	.203	.365	.383	.427	.054	.496	.077	.082	--	.001*	.117
Weikart													
Correlation	.1659	.0807	.1136	.2158	.1411	--	.4012	.0555	.1230	.2674	--	-.3030	--
Significance	.042*	.201	.119	.012*	.071	--	.001*	.282	.100	.002*	--	.001*	--

* $p < .05$ before correction for multiple comparisons.

[a] Code to subtests: 1 - information; 2 = similarities; 3 = arithmetic; 4 = vocabulary; 5 = comprehension; 6 = digit span; 7 = picture completion; 8 = picture arrangement; 9 = block design; 10 = object assembly; 11 = coding A; 12 = coding B; and 13 = mazes.

[b] Positive correlations indicate boys scored higher than girls. Negative correlations indicate girls scored higher than boys.

Table 24

Correlation of Mother's Education with WISC IQ Scores, by Project

Project (n)		Correlation	Significance (two-tailed)
Gordon (90)	FIQ[a]	.2960[d]	.004*
	VIQ[b]	.2982	.004*
	PIQ[c]	.2727	.007*
Gray (52)	FIQ	.0804	.288
	VIQ	.1979	.082
	PIQ	-.0419	.385
Levenstein (76)	FIQ	.1128	.166
	VIQ	.2173	.030*
	PIQ	-.0177	.440
Miller (141)	FIQ	.1951	.018*
	VIQ	.2191	.009*
	PIQ	.1178	.103
Palmer (132)	FIQ	.2666	.001*
	VIQ	.2686	.001*
	PIQ	.1644	.030*
Weikart (110)	FIQ	.2790	.002*
	VIQ	.2574	.003*
	PIQ	.2398	.006*

---

* $p < .05$ before correction for multiple comparisons.

[a] Full IQ score.

[b] Verbal IQ score.

[c] Performance IQ score.

[d] Positive correlations indicate boys scored higher than girls. Negative correlations indicate girls scored higher than boys.

Table 25

Correlation of Mother's Education with WISC Subtest Scores, by Project

	Subtest [a]												
Project	1	2	3	4	5	6	7	8	9	10	11	12	13
Gordon													
Correlation	.2614	.2147	.3777	.2085	.1766	--	.1845	.2090	.2042	.1391	--	.2571	--
Significance	.011*	.030*	.001*	.034*	.062	--	.054	.034*	.037*	.114	--	.012*	--
Gray													
Correlation	.0754	.1812	.1410	.1474	.2521	--	-.0198	-.0262	-.1282	-.0588	.0303	--	--
Significance	.300	.102	.162	.151	.037*	--	.445	.428	.185	.341	.416	--	--
Levenstein													
Correlation	.1664	-.1148	.1073	.3107	.1907	.2221	-.0022	-.0244	-.1352	-.0728	--	.1368	--
Significance	.075	.162	.178	.003*	.049*	.027	.493	.417	.122	.266	--	.119	--
Miller													
Correlation	.1093	.1481	.0422	.2533	.1558	.0071	.0072	.0703	.0151	.0138	--	.2239	.2090
Significance	.120	.055	.326	.003*	.047*	.470	.469	.226	.436	.441	--	.008*	.012*
Palmer													
Correlation	.2394	.2158	.1595	.2688	.2525	.2011	.1495	1118	.1245	.0051	--	.1244	.0747
Significance	.003*	.007*	.035 *	.001*	.002*	.011*	.044*	.103	.080	.477	--	.078	.198
Weikart													
Correlation	.2542	.1981	.2199	.1925	.0978	--	.1332	.0257	.1581	.0697	--	.2888	--
Significance	.004*	.019*	.010*	.022*	.155	--	.083	.395	.395	.050*	--	.001*	--

* p .05 before correction for multiple comparisons.

[a] Code to subtests: 1 = information; 2 = similarities; 3 = arithmetic; 4 = vocabulary; 5 = comprehension; 6 = digit span; 7 = picture completion; 8 = picture arrangement; 9 = block design; 10 = object assembly; 11 = coding A; 12 = coding B; and 13 = mazes.

Which Children Benefit Most on School Performance?

The reader will recall that earlier analyses concluded that children who participated in preschool programs were less likely to be assigned to special education classes or retained in grade later in their school careers. These were striking and important findings. However, it is possible that these results were obtained because preschool helped only the children who were already brighter or who came from somewhat more advantaged backgrounds. This hypothesis had been raised before (Bronfenbrenner, 1974) and there is some evidence to support it (e.g., Herzog, Newcomb & Cisin, 1974). Thus the next step is to consider what kinds of children were most affected by preschool.

Methods and Results

The possibility that preschool was differentially effective for different kinds of children was investigated first by determining whether the set of preschool-by-background crossproducts (interactions) would have a significant relationship to school outcomes when the set of background variables was controlled.[1] The background variables were: sex of child, ethnic background, family structure (father present or absent), family size (number of siblings), and maternal education. Standard regression analysis was employed. The interactions were represented by a set of multiplicative terms in the regression equation, formed by multiplying each background variable by the categorical preschool variable (preschool vs. control).[1] The increment in $\underline{R}^2$ due to addition of these non-linear crossproduct terms was then tested by the $\underline{F}$ statistic.[2]

---

[1] For a more detailed discussion, see J. Royce, 1979.

The results are shown in Table 26. The first column (R^2_A) shows the $\underline{R}^2$ for the set of linear terms (the background variables listed above). The second column (R^2_{AB}) indicates the total $\underline{R}^2$ for the set of linear terms and the set of interactive terms (set B).[3] As Table 26 shows, the $\underline{p}$ values range from .1921 to .6443. The set of interaction terms did not contribute significantly to the explained variation in school outcomes over and above that which was explained by the linear background terms alone.[4] Thus, we can conclude that preschool apparently helped children regardless of their sex, ethnic background and family background.

---

[1] Some data sets did not have a complete set of interactions. Palmer's all-male project did not have a sex-by-preschool crossproduct. The all-black samples (Gray, Weikart, Palmer) did not have ethnicity-by-preschool crossproducts. Gordon did not have a crossproduct for family structure.

[2] The $\underline{F}$ statistic is the ratio between the increment in $\underline{R}^2$ divided by the amount of residual variance after both the numerator and denominator have been divided by the appropriate degrees of freedom.

[3] The individual $\underline{F}$ values were converted to $\underline{z}^2$ and then summed. This sum was compared to a chi-square table with degrees of freedom equal to the number of combined $\underline{F}$'s. The last column presents the combined significance levels.

[4] Because the set of background interactions was not significant, the results from each background interaction were not pooled. The $\underline{F}$ values for these interactions (preschool by mother's educational level, father presence, number of siblings, sex, and ethnic group) are shown in Table 27A.

Another set of regression analyses was prepared to test whether preschool helped only the brightest children. In this case, the linear terms were preschool attendance and initial IQ scores[1] and the crossproduct was preschool-by-initial IQ.[2] Table 27 indicates that the interaction of initial IQ and preschool attendance was not significant; p values range from .3276 to .8291. Only two individual projects show a p value below .10. In Gray's group, treatment children with lower initial IQ scores were slightly more likely to avoid special education assignment than were treatment children with higher IQ scores (p = .0693).[3] The other project was Weikart's group, where treatment children with higher initial IQ scores were slightly less likely to be classified as underachievers compared to treatment children with lower initial IQ scores (p = .0731). Since Table 27 contains 17 p-values, one or two values around .06 or .07 are almost exactly what would be expected by chance.

The investigation was carried one step further by examining the number of positive versus the number of negative beta weights in order to determine whether there was a consistent pattern of children with initially low or high IQs benefitting more from preschool. No such pattern emerged. Thus, we can safely conclude that there is no evidence that preschool differentially benefitted the children who were already somewhat brighter. The same analyses seeking consistent patterns were performed for mother's education, father presence, number of siblings, sex, and ethnic background, with similar negative results (see Table 27A).

---

[1] Measured prior to beginning preschool.

[2] With only one crossproduct, the significance of the interaction was tested by testing the significance of the beta weight of the interaction. The beta weight was converted to a F ratio, as shown on Table 27.

[3] A negative regression coefficient indicates that children with lower initial IQs benefitted more from preschool.

In sum, these results imply that policy-makers need not worry about selecting which ethnic groups or family configurations or levels of intelligence to serve. All lower-income children can apparently benefit from preschool experience.

Table 26

The Relationship Between School Outcomes and the Set of Treatment by Background Crossproducts (interactions) when Linear Terms are Partialled Out

Data Set (n)	R^2_A	R^2_{AB}	F	p (2 tailed)	pooled z^2	Combined p (2 tailed)
Special Education Placement						
Experimental						
Gordon (64)	.0662 [a]	.1118 [b]	.941(3,55)[c]	.427		
Gray (51)	.2176	.2274	.130(4,41)	.971		
Weikart (123)	.1253	.1978	2.553(4,113)	.043		
Total Experimental (238)					4.737 [d]	.1921
Quasi-Experimental						
Beller (53)	.0920	.1269	.328(5,41)	.893		
Levenstein (123)	.0864	.1239	.950(5,111)	.452		
Miller (106)	.0568	.0784	.441(5,94)	.819		
Total all data sets (520)					5.373	.4970
Grade Retention						
Experimental						
Gordon (53)	.0942	.1559	1.072(3,44)	.371		
Gray (48)	.1271	.2560	1.646(4,38)	.183		
Palmer (219)	.0567	.0624	.428(3,211)	.733		
Weikart (97)	.0448	.0559	.256(4,87)	.906		
Total Experimental (417)					2.707	.6081
Quasi-Experimental						
Beller (53)	.2051	.2347	.317(5,41)	.900		
Levenstein (107)	.1213	.1990	1.843(5,95)	.112		
Miller (101)	.0397	.0455	.108(5,89)	.990		
Total all data sets (678)					5.249	.6296
Underachievement (Special Ed/Retention/Dropout)						
Experimental						
Gordon (64)	.0486	.0967	.976(3,55)	.411		
Gray (52)	.1610	.2736	1.628(4,42)	.185		
Palmer (219)	.0567	.0624	.428(3,211)	.733		
Weikart (123)	.1127	.1619	1.658(4,113)	.165		
Total Experimental (458)					4.480	.3449
Quasi-Experimental						
Beller (56)	.1771	.2090	.355(5,44)	.876		
Levenstein (125)	.0790	.1172	.978(5,113)	.435		
Miller (120)	.0438	.0574	.312(5,108)	.905		
Total all data sets (759)					5.128	.6443

Note. Equation: ACH = MED + FP + SIBS + SEX + ETH + PC + PCMED + PCFP + PCSIBS + PCSEX + PCETH. Legend: PCMED = preschool x mother's educational level; PCFP = preschool x father present; PCSIBS = preschool x siblings; PCSEX = preschool x sex; PCETH = preschool x ethnicity; also see Table 12.

[a] R^2_A = step 1, linear terms only.

[b] R^2_{AB} = total R^2 = linear terms and interactions.

[c] F = ratio between increment in R^2 divided by amount residual variance (after dividing numerator and denominator by appropriate degrees of freedom). Numbers in parentheses indicate degrees of freedom.

[d] $z^2 = X^2$; degrees of freedom = number of pooled F's.

Table 27

The Effect of Initial IQ Level on the Relationship Between Early Education and School Outcomes

Data Set	*n*	PREIQ x PC Coeff.	*F*	Signif. Level (2 tailed)	Pooled *z* score	Pooled *p* (2 tailed)
			Special Ed Placement			
Approx. Experimental Design						
Gordon	43	.018 [a]	1.331[b]	.2585		
Gray	49	-.012	3.499	.0693		
Weikart	123	.018	2.386	.1258		
Experimental Total	215				.4874[c]	.6260
Quasi-Experimental						
Beller	65	.002	.273	.6047		
Levenstein	121	.007	.692	.4082		
All Data Sets					.9790	.3276
			Grade Retention			
Approx. Experimental Design						
Gordon	32	-.021	1.118	.3034		
Gray	46	-.010	.757	.3919		
Palmer	134	-.004	.441	.5086		
Weikart	97	.006	.400	.5297		
Experimental Total	309				-.9590	(.3376)[d]
Quasi-Experimental						
Beller	65	.00007	.000	1.0000		
Levenstein	106	.011	1.968	.1647		
All Data Sets					-.2158	(.8291)[d]
			Underachievement (Special Ed/Retention/Dropout)			
Approx. Experimental Design						
Gordon	43	.008	.187	.6697		
Gray	51	-.010	1.009	.3228		
Palmer	134	-.004	.441	.5086		
Weikart	123	.023	3.282	.0731		
Experimental Total	351				.2843	.7762
Quasi-Experimental						
Beller	68	.004	.191	.6648		
Levenstein	123	.011	1.480	.2271		
All Data Sets					.9021	.3670

Note. Equation: ACH = PC + IQPRE + PCIQPRE. Legend: PCIQPRE = preschool x initial IQ score. See Table 12 for legend.

[a] Unstandardized regression coefficient for initial IQ x preschool interaction. Positive coefficient indicates benefit favors children with higher initial IQ score.

[b] *F* test for significance of interaction beta weight. Simultaneous (standard regression) method with interaction term added after all other variables.

[c] Pooled $z = \Sigma z_i / \sqrt{k}$.

[d] Parentheses indicate negative pooled *z* score. That is, benefit favors children with lower initial IQ score.

Table 27A

F Values for Beta Weight of Interaction Terms (Crossproducts of Treatment and Background Characteristics)

Data Set (n)	PCMED	PCFP	PCSIB	PCSEX	PCETH
		Special Education Placement			
Experimental					
Gordon (64)	.337		- .790	- .336	
Gray (51)	.066	.000	- .263	- .297	
Weikart (123)	.021	-6.799**	-5.326*	- .573	
Quasi-Experimental					
Beller (53)	.068	.045	-1.45	- .011	- .261
Levenstein (123)	.000	- .425	.149	3.042	.121
Miller (106)	.772	.124	- .000	- .007	- .767
No. of positive beta weights:[a]	5	2	1	1	1
		Grade Retention			
Experimental					
Gordon (53)	1.535		- .005	-1.044	
Gray (48)	.194	-2.684	.051	-3.267	
Palmer (219)	- .000	1.068	- .251		
Weikart (97)	.042	- .132	.117	.433	
Quasi-Experimental					
Beller (53)	- .109	.005	.038	.060	- .143
Levenstein (107)	-3.065	2.184	1.781	.038	-8.024**
Miller (101)	- .017	- .188	.083	.007	- .074
No. of positive beta weights:	3	3	5	4	0
		Underachievement (Special Ed/Retention/Dropout)			
Experimental					
Gordon (64)	.990		- .326	- .327	
Gray (52)	.403	-2.523	- .040	-3.667	
Palmer (219)	- .000	1.068	- .251		
Weikart (123)	.143	-5.248*	-2.508	- .027	
Quasi-Experimental					
Beller (56)	- .106	.134	- .004	.015	- .284
Levenstein (125)	.009	- .439	.460	1.742	-1.938
Miller (120)	.464	.104	.004	- .008	- .706
No. of positive beta weights:	5	3	2	2	0

Note. Sign of F is sign of beta weight for interaction term. Positive sign indicates higher level of mother's education; father present; fewer siblings; male student and black ethnic group.
Legend: PCMED = preschool x mother's education; PCFP = preschool x father presence; PCSIBS = preschool x number of siblings; PCSEX = preschool x sex; PCETH = preschool x ethnicity.

[a] These totals exclude beta weights of .000.

* p = <.05

** p = <.01.

What Kinds of Programs Were Most Effective?

In considering differential effectiveness of preschool programs, the question arises: what kinds of programs were most effective? The answer to this question is especially important to policy-makers. And it is a question which has been difficult to answer. For example, two very recent early intervention programs report considerable success in raising the IQ scores of low-income Black infants (Garber & Heber, 1977) and under-nourished Colombian children (McKay, et al., 1978), but neither was able to stipulate what parts of the program were responsible for the gains. The preliminary report of the National Day Care Study (Abt, 1978) has been somewhat more successful; early results point to the importance of group size and teacher training in child-related fields as important influences on both classroom behavior and cognitive outcomes of participating children.

Since some of the Consortium's principal investigators had built program variations into their original designs, we undertook analysis of the Consortium data in the hopes that it could shed some light on the important question of program effectiveness. The reader should bear in mind, however, that this is a secondary analysis of data which were not, in many cases, originally designed to answer this particular question. Specifically, problems arose in choosing a common measure of program effectiveness and, consequently, in comparing programs across sites (e.g., comparing

Miller's programs to Karnes' programs), as will be detailed below. The current report should be regarded as one more step in our continuing search for meaningful and valid ways of assessing the program effectiveness of the Consortium projects.

Vopava and Royce (1978) addressed the question of what kinds of Consortium programs were most effective in detail.[1] The set of children within a project who received the same type of curriculum (e.g., Montessori, Bereiter-Engelmann, etc.) was the unit of analysis. There were 21 such subgroups. Twenty-four program characteristics -- such as age of child at entry, presence or absence of home visits, amount of teaching structure, etc. -- were assessed, using reduction in special education placements in later school years as the measure of effectiveness.[2] Nine program characteristics were significantly related to effectiveness.[3] Five of those program characteristics were highly intercorrelated: age of entry, home visits, program goals for parents, parental involvement and number of children per adult. Together, they suggest that the most effective programs involve one instructor working with an infant or toddler and his/her parent in the home. Because of the high intercorrelations, however, it was not possible to pinpoint the relative importance of the five variables in reducing later placements in special education.

---

[1] See also Consortium (1977) for another, earlier effort.

[2] See the methods section of this chapter for a full description of how this measure is derived.

[3] They were: age of entry (r = -.64), program goals for parents (r = .83), home visits (r = .64), center-based program (r = -.56), children per adult (r = .83), professional staff (r = -.50), parental involvement (r = .60), hours per year of program (r = .73), adult contact hours (r = -.53).

We have reviewed the above results in accordance with our general policy of reporting results reasonably promptly but then challenging them by testing alternative hypotheses.[1] In searching for the most effective kinds of preschool programs, our work this year had the following aims: (1) to increase the power of the statistical analyses by reducing the number of program characteristics, primarily by combining those characteristics which were highly intercorrelated; and (2) to test the robustness of any positive findings. In brief preview of the findings, we have had to consider the results as inconclusive. Using the independent and dependent variables,described below, we were unable to reject the null hypothesis that variations in program characteristics are unrelated to program effectiveness.

Methods. For all the analyses in this section, the unit of analysis was the subgroup within a project, defined as the set of children who received the same type of curriculum (e.g.,

---

[1] This policy served us well in reporting the findings on the overall effects of preschool on later school performance, Our early analyses found positive effects (Consortium, 1977). The findings then survived the numerous challenges documented in this report, serving to increase our confidence in those findings.

Montessori, DARCEE, etc.).[1] There were 21 such curriculum subgroups across different projects, as shown in Table 28.

The independent variables -- program characteristics -- for these analyses are shown in Table 29. These are based on ratings submitted by the principal investigators. Some were quite objective, such as the average age at which children entered the program. Others were more subjective, such as the importance of language goals. Earlier reports and papers from this research group (Consortium, 1977; Vopava & Royce, 1978) reported analyses on 24 such program characteristics. We have since pruned this number down to 10. Some of the 24 variables were eliminated because they were conceptually redundant with variables left in, and others were eliminated because they correlated very highly with the remaining variables. This pruning process is described in more detail by Ypelaar (1978).

For the dependent variable, the reduction in special education placements in the public schools was used as the measure of program effectiveness for these analyses.[2] This reduction was defined as the percentage of experimental-group children not placed in special education classes, minus the comparable percentage for the control group from the same project. Analysis of covariance was used to correct for any differences between experimental and control groups on pretest IQ scores. Thus, in this section, the measure of program effectiveness is the proportion of a program's children not placed in special education, minus the same proportion for a matched control group, after adjusting for any pretest IQ differences between the program and control groups. The values found in this way are shown in Table 28.[1]

---

[1] The individual child cannot be the unit of analysis, since a single exceptionally effective teacher, or some other single chance effect, might well improve the performance of all the children in a group. Thus the children in a subgroup cannot be considered to be statistically independent of each other.

[2] Since the Palmer project had no data on special education placements, it was not used in these analyses.

Particular difficulties arose in the analysis of three projects using this measure of effectiveness. Karnes had no control groups, so the procedure could not be applied to her data. Instead we compared the proportions in each subgroup to the proportions in the five Karnes subgroups combined. This would tend to make the Karnes subgroups look less effective than those of other projects since presumably all her program subgroups had fewer children placed in special education than a control group would have. Miller's control group came from a higher income level than her experimental groups and had significantly higher percentages of white and father-present families. As a result, the measures of program effectiveness for her data contain a strong conservative bias. Unfortunately, a conservative bias in the analysis of overall program effectiveness can introduce a liberal bias in the present analysis. For instance, all of Miller's programs used older children; thus, underestimating the effectiveness of Miller's programs will make programs for young children look good in comparison to programs for older children, thereby introducing a spurious correlation between age and program effectiveness. Beller's data was similarly over-conservative because the Philadelphia school system placed almost no children into special education. Thus, even if Beller's project had been totally effective in eliminating special education placements, Beller's project would have appeared to be below average in effectiveness by this measure of effectiveness. We considered other measures of effectiveness, such as the ratio between program-group special education placements and control-group special education placements, but all were rejected for technical reasons, even though they might have solved this particular problem for Beller's data.[1]

---

[1] We did not adjust the subgroups from Gordon's project for pretest IQ since most of the children in his project had not been administered an IQ test before they began the program. For the few children who had been pretested, control children had a higher mean pretest IQ than program children; thus the bias appears to be in the conservative direction. Miller's subgroups were not adjusted for pretest IQ since the first IQ test she administered was 6-8 weeks after the program had started.

The variable of parental involvement concretely illustrates this problem. As noted above, the Vopava and Royce (1978) analysis suggested that five intercorrelated variables caused effective outcomes. In the current analyses, four of those variables -- goals for parents, home visits, parental involvement, and children per adult -- have been combined with child group size into a single variable, labelled "Parental Involvement." Looking at the subgroups' ratings on this variable, however, it turned out that the Karnes, Miller, and Beller projects accounted for all the subgroups with low or moderate parental involvement. Since all these programs had over-conservative estimates of effectiveness, analysis would result in a spurious correlation between parental involvement and program effectiveness. If the three projects were deleted, there would be no variance on parental involvement -- all remaining programs had high parental involvement. We are thus particularly unable to assess the importance of parental involvement in the preschool program.

---

[1] Two different approaches, to be described later, were used to circumvent these difficulties with the Karnes, Miller, and Beller data.

Because our sample consisted of only 21 observations (subgroups), and because we cannot be sure of normal distributions, the jackknife technique was used for estimates and significance tests. This technique is powerful for small non-normal samples. In the jackknife technique[1] one observation, or a group of observations, is removed from the sample one at a time and the statistic of interest is computed. Then the observation is replaced, another observation is removed, and the statistic is recomputed. This procedure is repeated until all observations have been removed once. The statistics computed in this way are used to compute both an estimate of the statistic of interest and a standard error of that estimate. The ratio of the estimate to its standard error can then be tested as a _t_ statistic in the usual way.[2] The jackknife technique is so named because it is so versatile -- significance tests can be performed on virtually any statistic by this technique. In addition, since observations are systematically deleted, it serves as a test of robustness.

---

[1] For a fuller description of the technique, see Mosteller and Tukey (1977).

[2] See Ypelaar (1978) for more details.

Results

Four different jackknife analyses were performed; only the third and fourth will be reported in detail here. The first two jackknife analyses led to the discovery of the problems with the Beller, Karnes, and Miller projects detailed above.

In the third jackknife analysis, we excluded the three problematic projects entirely from the analysis and then removed the remaining subgroups (N = 10) one by one. Columns 1 and 2 of Table 30 show the results. One variable -- preservice training of teachers -- was significantly related to effectiveness. An examination of the raw data, however, indicated that this finding was due to only one project: Weikart.[1] A finding based on only one project is inconclusive because, since the location of the project and the project curricula were confounded, it might be an artifact of the policies of that particular school district.

In the fourth jackknife analysis we looked only at intra-project differences. This procedure was followed because it was virtually impossible to compare the average effectiveness of the Karnes, Miller, or Beller programs to the average effectiveness of the other preschool programs, as explained earlier. Imagine that in every project which included separate programs for young children and older children, the programs for younger children had shown up as more effective. This imaginary result would be extremely interesting and would not be hampered by the difficulties of comparing effectiveness across projects. To search for results like this, we examined only intra-project differences in effectiveness between programs by adjusting the average effectiveness measure for each project to zero. Thus a negative effectiveness by this measure would mean only that a program was less effective than the average of the programs studied by that particular investigator. As Table 30, columns three and four show, this jackknife analysis found no significant correlations between program effectiveness and program characteristics. Since there were no positive findings in this analysis, we have to conclude that no one type of program can confidently be identified as more effective than any other.

---

[1] Staff training implies an existing curriculum in which staff are trained. For the Perry Preschool the process might better be characterized as program development with input from both classroom and non-classroom staff throughout the year.

In review, we have looked at the data in three different ways -- once in Consortium (1977), once by Vopava and Royce (1978), and once again in this chapter. However, all of these analyses have difficulties. Our first analysis (Consortium, 1977) suffered from using the child rather than the subgroup as the unit of analysis, while the second (Vopava and Royce, 1978) underestimated the difficulties involved in using the Miller and Beller data on reductions in special education placement. The present analyses suffer from several difficulties outlined in the next paragraph.

How can we interpret the current results? We must emphasize that we have not shown that there are no differences in program effectiveness. For example, parental involvement may play a crucial role in preschool education, but our data cannot address that issue. One reason for the lack of findings is that the statistical technique we used is very conservative, leading to a loss of power. When we have so few observations to begin with, this loss of power can be ill-afforded. A second reason is that the dependent variable used here (frequency of placement in special education) may not be the most sensitive variable for differentiating among different programs at different sites; frequency of special education placements varies randomly from city to city and random error lowers the power of the analysis. Third, all of the Consortium preschool programs were exceptionally well-run programs. It is more difficult to find differences among programs which were uniformly well-run than among programs which were not. Fourth, we must remember the general statistical principle that when the null hypothesis fails to be rejected, the null hypothesis is not therefore proved.

The Consortium is currently planning to reexamine its data in still other ways which may yield different results. We plan to examine program characteristics in depth in three of the projects which had planned curriculum variations (Karnes, Miller, Weikart). We plan to use achievement test scores as a dependent variable. We will attempt to use alternative methods of analysis which may be more powerful than the statistical analyses reported here. In the meantime, we would caution against putting too much reliance on either the findings reported earlier or the lack of findings reported here. Instead, we note that the data suggest the real need for new experiments specifically designed to separate and measure the effects of these important program variables.

Table 28

Subgroups and Effectiveness Scores[1] in Reducing Special Ed Placements

Project	Subgroup	Adjusted Difference
Beller T_1 *	Nursery School	0.007
Gordon $T_{1,3,4}$	One year home visits; 2nd year home visits and home learning center	0.234
$T_{2,5,6}$	One or two years of home visits only	0.371
T_7	One year of home visits and home learning center	0.332
Gray Early Training Project T_1	Two year program	0.250
T_2	One year program	0.266
Karnes T_1	Traditional	0.002
T_2	Bereiter-Engelmann	-0.139
T_3	GOAL	0.077
T_6	Montessori	0.159
T_7	Community Integrated	-0.101
Levenstein $T_{1,16}$	One year program	0.257
$T_{5,7,8}$	Two year program	0.203
$T_{14,15}$	One full year and one short year program	0.383
Miller T_9	DARCEE	0.025
T_1	DARCEE with home visits	-0.250
T_2	Bereiter Engelmann	-0.106
T_3	Montessori	-0.051
T_4	Traditional	0.000
Weikart/Perry Preschool T_1	Two years	0.129
T_2	One year	0.124

---

[1] Source: Vopava and Royce (1978), p. 14.

* T refers to treatment group. Each number indicates a different program approach within the project as a whole.

Table 29

Reduced Set of Program Variables Included in the Analysis

Age of child at entry

Length of program in years

Months per year that program operated

Parental Involvement[1]

Center-based program

Professional vs. paraprofessional Staff[2]

Preservice training for staff

Language goals for children

Amount of teaching structure

Hours per year of program[3]

---

[1] This represents a new variable, created by combining the following five variables: (1) goals for parents; (2) parental involvement; (3) home visits; (4) children/adult; (5) child group size. These five variables were so highly intercorrelated as to represent a single variable. See Ypelaar (1978)

[2] This represents a new variable, created by combining the following three dichotomous variables: (1) use of professional staff; (2) use of paraprofessional staff; (3) use of volunteer staff. The new variable was coded dichotomously, representing the highest level of professionalism attained (professional vs. paraprofessional). See Ypelaar (1978)

[3] Transformed to log (hours) to make the variable more normally distributed for correlational analysis. See Ypelaar (1978)

Table 30

Jacknifed Estimates of the Effectiveness Correlation and Associated Probability Level for Each Program Variable

	Between-Projects Analysis		Within-Projects Analysis	
	Subgroup r*+	Corrected Signif. Level	Subgroup r*	Corrected Signif. Level
Age	-.234	--	.022	--
Years of Program	-.336	--	-.135	--
Month/Year	.269	--	‡	‡
Parental Involvement	‡	‡	-.213	--
Center	-.469	--	-.023	--
Paraprofessional vs. Professional Staff	-.277	--	.154	--
Preservice Training	.806	.01	‡	‡
Language Goals	-.059	--	-.255	--
Teaching Structure	.256	--	-.082	--
LN (Hours/yr.)	-.669	--	-.084	--

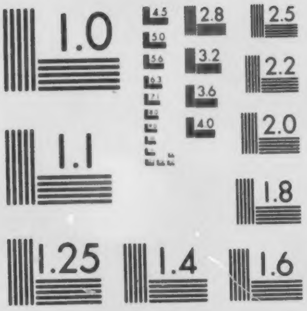
Note. All significance levels are two-tailed. (--) indicates probability larger than .05.

* Refers to jackknifed r.

+ Without Beller, Karnes and Miller projects.

‡ No variability.

LB


1.0
4.5
5.0
5.6
6.3
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Matching Children to Programs

The failure to find evidence in our data of differential program effectiveness was due in some part to problems with control groups in the Beller, Karnes, and Miller projects. Hence we were forced to eliminate these projects from the analysis or to perform within-project analyses only. However, another question of interest concerning program characteristics arises which may be answered without including control groups in the analysis. In the groups of treatment children, did some kinds of children respond more to certain programs than to others? For example, did children from large families benefit more from programs with high structure than did children from small families? We can answer these questions by examining statistical interactions. The question just posed, for example, can be examined by studying the interaction between family size and program structure. If such effects were found to exist, it would be possible to make specific recommendations in matching children to programs.

Methods

To assess the presence or absence of interaction effects, regression analyses using second order terms were performed. The measures of effectiveness of programs -- the dependent variables -- were placement in special education and retention in grade. They were chosen because earlier analyses had indicated that preschool attendance affected them and because they represent important, fiscally relevant outcomes for individual children. The program variables used included eight of the ten variables which were least collinear: age of the subjects at onset of intervention, the length of intervention, the number of months of intervention, the number of hours per year of intervention, the level of parental involvement, the presence or absence of language goals, preservice training for program teachers, and the degree of structure in the teaching methods. These variables represent only a subset, although a select subset, of all possible program characteristics.

The child characteristics included the child's pretest IQ score and birth order, level of mother's education, family size (number of siblings), family structure (father presence or absence) and the initial hopes of the mother for the educational attainment of the child. The latter variable was coded from the intake interview before preschool.

All treatment children from the following projects were included in the analyses: Beller, Gordon, Gray, Karnes, Levenstein, Miller, Palmer Weikart.

The regression analysis proceeded as follows. First, all linear terms (i.e. the eight program variables and the six child characteristics) were adjusted to means of zero. Then 8 x 6 or 48 interaction terms were created by multiplying each of the program variables by each of the child variables. To avoid excessive numbers of regression terms in any one regression equation, we studied only the eight interaction terms pertaining to a single child characteristic at any one time (that is, in any one regression). A forced-order regression was performed in two steps, entering first the eight program variables and the one selected child characteristic, and second entering the eight interaction terms pertaining to that characteristic. The increase in $\underline{R}^2$ obtained in this second step (in other words, the proportion of variance due to interaction) was then tested for significance. This procedure was repeated for each child characteristic variable. This whole process was performed for two dependent variables: retention in grade, and assignment to special education. Since we were studying here the relative effectiveness of various treatments, only children who had attended infant or preschool programs were used in these analyses.

Results

Table 31 shows the results of the analyses with assignment to special education as the measure of program effectiveness.[1] Clearly, there were no significant interactions involving the child characteristics and program characteristics. While pretest IQ score and maternal education approach significance, they would not survive the procedure of correcting for multiple comparisons. Thus we cannot even consider them as trends.[2] Similarly, Table 32 reveals no significant interaction effects when retention in grade is the measure of effectiveness.

---

[1] The numbers reported represent the $\underline{F}$ values and $\underline{p}$ values associated with the increase in $\underline{R}^2$ after inclusion of the interaction terms on the second step of the regression.

[2] If any of the variables had proven significant, the next step would be to examine individual program characteristics to determine which were most effective.

In sum, there appear to be no systematic benefits derived by matching certain kinds of children to certain kinds of programs, at least insofar as we have been able to measure using molar measures of program effectiveness. This result is discouraging in the sense that we can provide no specific guidance to policy-makers who design programs. It is important to point out, however, that these programs were well-operated. In that sense, we may say that infant and preschool programs, if they are well-run, benefit all kinds of lower-income children.

Table 31

F Values and p Values for Incremental R^2 of Interaction Terms[a] Using Special Ed as Dependent Variable

Child Characteristic	F	p	N[b]
Pretest IQ[c]	2.0651	.0747	128
Mother's Education	1.68842	.0996	394
Father Presence	.81653	.5884	377
Birth Order	.84316	.5650	418
Number of Siblings	.61224	.7677	418
Mother's Pre-intervention Hopes for Child Education	2.117	.1298	63

[a] This table indicates the extent to which the interaction terms significantly increased our ability to predict the dependent variables.

[b] Size of sample varies because some subjects had missing data on some characteristics.

[c] Only those projects with Stanford-Binet IQ scores were included in the analyses: Beller, Gray, and Weikart.

Table 32

F Values and p Values for Incremental R^2 of Interaction Terms[a] Using Retention as Dependent Variable

Child Characteristic	F	p	N[b]
Pretest IQ[c]	.6769	.6915	208
Mother's Education	.63483	.7486	519
Father Presence	.60428	.7745	507
Birth Order	1.13953	.3348	532
Number of Siblings	1.32164	.2299	418
Mother's Pre-intervention Hopes for Child Education	1.48609	.2371	51

[a] This table indicates the extent to which the interaction terms significantly increased our ability to predict the dependent variables.

[b] Size of sample varies because some subjects had missing data on some characteristics.

[c] Only those projects with Stanford-Binet IQ scores were included in the analyses: Beller, Gray, Palmer, and Weikart.

Summary of Differential Effects

The first section ("Which Children Benefit Most on WISC-R Test Scores?") reported no differential gains on later intelligence tests for boys as opposed to girls or for children whose mothers had more vs. less education.

The hypothesis tested in the section entitled, "Which Children Benefit Most on School Performance?", was: the effects of preschool on later school outcome are due to the fact that preschool benefits some kinds of children more than others. The hypothesis was not confirmed. There was no evidence in these data that the preschool experience helped boys more or less than girls, white children more or less than Blacks, children with higher initial IQ scores more or less than those with lower initial IQ scores, children from two-parent families more or less than those from single-parent families, or children with fewer siblings more or less than children with many siblings.

In the section entitled, "What Kinds of Programs Were Most Effective?", we attempted to answer the question: what kinds of preschool programs were most effective? Ten different variables were used (as independent variables) to characterize the programs. The measure of effectiveness was the reduction in special education placements associated with preschool attendance. In general, we were unable to show that some program characteristics were superior to others in reducing placements in special education.

In the section entitled, "Matching Children to Programs", we attempted to determine whether the various program characteristics helped some kinds of children more than others. Avoiding assignment to special education classes and retention in grade were used as measures of benefits to children. Once again, there was no evidence that some of the children who attended preschool derived more benefit from certain kinds of programs than others.

Taken together, all four results suggest that all low-income children can benefit from attending high quality programs.

THE EFFECTS OF FAMILY CHARACTERISTICS ON SCHOOL OUTCOMES FOR LOW-INCOME CHILDREN

Predicting School Outcomes

Throughout this report we have concentrated our efforts on comparing children with preschool experience to children who lacked that experience. However, these "treatment" and "control" children were originally chosen because they represented a population at risk for school failure: low-income, primarily Black children. The Consortium data bank represents a rich source of data concerning this population, containing background information, IQ scores and school outcome measures for more than 3,000 children over a period of many years. In order to provide a general context for the treatment/control differences reported throughout this report, we partialled out the effects due to preschool and investigated the children's school performances in light of a number of characteristics measured prior to enrollment in preschool. In effect, we attempted to assess how well these measures predicted the later school performance of this sample of low-income children. Included in the analyses as predictors were family structure (father present vs. absent), number of siblings, mother's level of education, child's sex, and child's initial IQ scores and IQ at age 6.

Methods

As in earlier analyses, the variables used to assess school performance were assignment to special education classes, retention in grade, and underachievement (a composite variable consisting of special education placement and/or retention in grade and/or dropping out of school). Regression analysis was employed to allow simultaneous consideration of the effects of the various child characteristics and family background variables while partialling out the effects of attending preschool.

Results

The first such analysis included the following independent variables: child's sex, ethnic background, IQ score at age 6, family structure (father present vs. absent), family size (number of siblings) and mother's level of education. We tested the hypothesis that mother's educational level would predict later school outcomes when the other background variables, preschool attendance, and IQ score at age 6 were controlled. As Table 33 shows, there appears to be no such effect for assignment to special education ($p = .4239$, $N = 557$) or retention in grade ($p = .5077$, $N = 678$). For the composite underachievement variable, results are marginally significant ($p = .0673$, $N = 774$) when pooled across all projects.

The analyses were repeated with the difference that IQ score at age 6 was not partialled out, and a somewhat different picture emerged. As shown in Table 34, mother's educational level predicted all three school performance measures: assignment to special education (pooled $p = .0233$, $N = 588$), grade retention (pooled $p = .0215$, $N = 723$), and underachievement (pooled $p = .0003$, $N = 820$). The findings were not robust for either special education placement or retention in grade. For underachievement, however, the results remained significant (p = .020) even when the two data sets with the strongest findings (Beller and Weikart) were deleted.

In brief, level of maternal education strongly and robustly predicted children's underachievement when effects of other background characteristics and effects of preschool attendance were partialled out. The results were significant but not robust for the school performance measures of assignment to special education and retention in grade. However, when the effects of the child's IQ score at age 6 on later school performance were partialled out, maternal education was not related to later school performance.

How well did the other background variables predict later school performance? Regression analysis was again employed with the set of independent variables of sex, ethnic background, family structure (father presence) and family size and partialling out the effects of maternal education and preschool attendance. As Table 35 shows, the set of independent variables was not associated with any of the school outcome measures. Apparently, then, only maternal education, of all the background variables gathered before preschool programs commenced, significantly predicted children's later school performance for this sample of children (and only when its effect on IQ score at age 6 was not partialled out of the equation).

Table 33

The Effect of Mother's Education on School Outcomes when Preschool Attendance, Other Background Variables, and IQ at Age 6 Are Controlled

Data Set	n	Mother's Educ. Coeff.	F	p (2 tailed)	Pooled z score	Mean p (2 tailed)
			Special Ed Placement			
Gordon	62	-.0260[a]	.725[b]	(.4003)[c]		
Gray	50	.0213	1.760	.1941		
Weikart	120	.0249	2.346	.1292		
Beller	50	-.0189	.679	(.4173)		
Levenstein	116	-.0263	1.964	(.1649)		
Miller	106	.0378	3.088	.0827		
Karnes	53	.0262	.377	.5444		
All Data Sets	557				.7997[d]	.4239
			Grade Retention			
Gordon	51	-.04[illegible]8	1.746	(.1956)		
Gray	46	.0264	.843	.3671		
Palmer	195	.0321	3.054	.0826		
Weikart	94	-.0056	.116	(.7350)		
Beller	50	.0723	3.207	.0822		
Levenstein	101	-.0106	.360	(.5510)		
Miller	101	-.0082	.337	(.5640)		
Karnes	40	.0123	.095	.7616		
All Data Sets	678				.6624	.5077
		Underachievement (Special Ed/Retention/Dropout)				
Gordon	62	-.0415	1.530	(.2234)		
Gray	50	.0315	1.635	.2104		
Palmer	195	.0321	3.054	.0826		
Weikart	120	.0175	.970	.3278		
Beller	53	.0656	2.735	.1069		
Levenstein	118	-.0226	1.125	(.2922)		
Miller	120	.0291	2.110	.1500		
Karnes	56	.0181	.187	.6689		
All Data Sets	774				1.8298	.0673

Note. Equation: ACH = IQ6 + PC + FP + SIBS + SEX + ETH + MED. For legend, see Table 12.

[a] Unstandardized regression coefficient for mother's educational level.

[b] F test for significance of mother's education coefficient.

[c] Figures in parentheses are in reverse direction.

[d] Pooled $z = \Sigma z_i / \sqrt{k}$.

Table 34

The Effect of Mother's Education on School Outcomes when Preschool Attendance and Other Background Variables Are Controlled

Data Set	n	Mother's Educ. Coeff.	F	p (2 tailed)	Pooled z score	Mean p (2 tailed)
Special Ed Placement						
Gordon	64	-.001[a]	.002[b]	(.9647)[c]		
Gray	51	.024	2.004	.1660		
Weikart	123	.043	6.720	.0109		
Beller	53	-.006	.069	(.7950)		
Levenstein	123	-.019	.986	(.3238)		
Miller	106	.043	3.863	.0527		
Karnes	58	.056	2.103	.1549		
All Data Sets	578				2.2677[d]	.0233
Grade Retention						
Gordon	53	-.015	.141	(.7104)		
Gray	48	.038	1.650	.2085		
Palmer	219	.037	4.108	.0442		
Weikart	97	.001	.002	.9645		
Beller	53	.108	6.748	.0130		
Levenstein	107	-.004	.067	(.7968)		
Miller	101	-.004	.084	(.7732)		
Karnes	45	.058	2.785	.1052		
All Data Sets	723				2.2988	.0215
Underachievement (Special Ed/Retention/Dropout)						
Gordon	64	.001	.000			
Gray	52	.041	2.479	.1241		
Palmer	219	.037	4.108	.0442		
Weikart	123	.040	4.693	.0327		
Beller	56	.110	6.764	.0127		
Levenstein	125	-.014	.452	(.5036)		
Miller	120	.034	2.894	.0924		
Karnes	61	.043	1.287	.2636		
All Data Sets	820				3.6458	.0003

Note. Equation: ACH = FP + SIBS + SEX + ETH + PC + MED. For legend, see Table 12.

[a] Unstandardized regression coefficient for mother's educational level.

[b] F test for significance of mother's education coefficient.

[c] Figures in parentheses are in reverse direction.

[d] Pooled $z = \Sigma z_i / \sqrt{k}$.

Table 35

The Effect of Other Background Variables on School Outcomes, When Mother's Education and Preschool Attendance Are Controlled

Data Set	n	R^2_A	R^2_{AB}	F	p (2 tailed)	Pooled z^2	Pooled p (2 tailed)
Special Ed Placement							
Gordon	64	.0221 [a]	.0662 [b]	.913 (3,58) [c]	.4404		
Gray	51	.2012	.2176	.314 (3,45)	.8149		
Weikart	123	.0923	.1253	1.471 (3,117)	.2260		
Beller	53	.0017	.0920	1.144 (4,46)	.3480		
Levenstein	123	.0721	.0864	.454 (4,116)	.7694		
Miller	106	.0432	.0568	.357 (4,99)	.8387		
Karnes	58	.0099	.0803	.995 (4,52)	.4186		
All Data Sets	578					3.7773 [d]	.8050
Grade Retention							
Gordon	53	.0037	.0942	1.565 (3,47)	.2106		
Gray	48	.0702	.1271	.913 (3,42)	.4431		
Palmer	219	.0537	.0567	.340 (2,214)	.7119		
Weikart	97	.0355	.0448	.295 (3,91)	.8287		
Beller	53	.1617	.2051	.628 (4,46)	.6450		
Levenstein	107	.0098	.1213	3.172 (4,100)	.0169		
Miller	101	.0170	.0397	.556 (4,94)	.6955		
Karnes	45	.0346	.0976	.681 (4,39)	.6095		
All Data Sets	723					8.6830	.3697
Underachievement (Special Ed/Retention/Dropout)							
Gordon	64	.0092	.0486	.801 (3,53)	.4986		
Gray	52	.0987	.1610	1.139 (3,46)	.3435		
Palmer	219	.0537	.0567	.340 (2,214)	.7119		
Weikart	123	.0951	.1127	.774 (3,117)	.5110		
Beller	56	.1479	.1771	.435 (4,49)	.7829		
Levenstein	125	.0340	.0790	1.441 (4,118)	.2246		
Miller	120	.0279	.0438	.956 (4,113)	.4345		
Karnes	61	.0031	.0843	1.219 (4,55)	.3132		
All Data Sets	820					5.1009	.7467

Note. Equation: ACH = MED + PC + FP + SIBS + SEX + ETH. For legend, see Table 12.

[a] R^2_A = first step (MED and PC only) stepwise regression

[b] total R^2

[c] F = ratio between increment in R^2 divided by amount of residual variance (after dividing numerator and denominator by appropriate degrees of freedom). Numbers in parentheses indicate degrees of freedom.

[d] $\Sigma z^2 = x^2$; degrees of freedom = number of pooled F's.

REVIEW OF THE DETERMINANTS OF SPECIAL EDUCATION PLACEMENTS

We are now in a position to integrate some of our findings and attempt to pinpoint the relative influence of the preschool experience on the development of this sample of lower-income children. We should point out that the discussion which follows is based on longitudinal data, that is, information about each child at several specific times in his/her life span.[1] The details of family structure and size, maternal education, and initial IQ score were collected before the treatment children were enrolled in the preschool programs. These variables provide a picture, albeit limited, of the child's circumstances before experiencing intervention. The age at which the children began attending the programs is known. Most subjects were given a posttest IQ test when they were 6 years old,[2] at approximately the age when most children enter first grade. School records for the intervening 3 to 13 years provide information about the child's school performance up to the time of the 1976-77 follow-up.

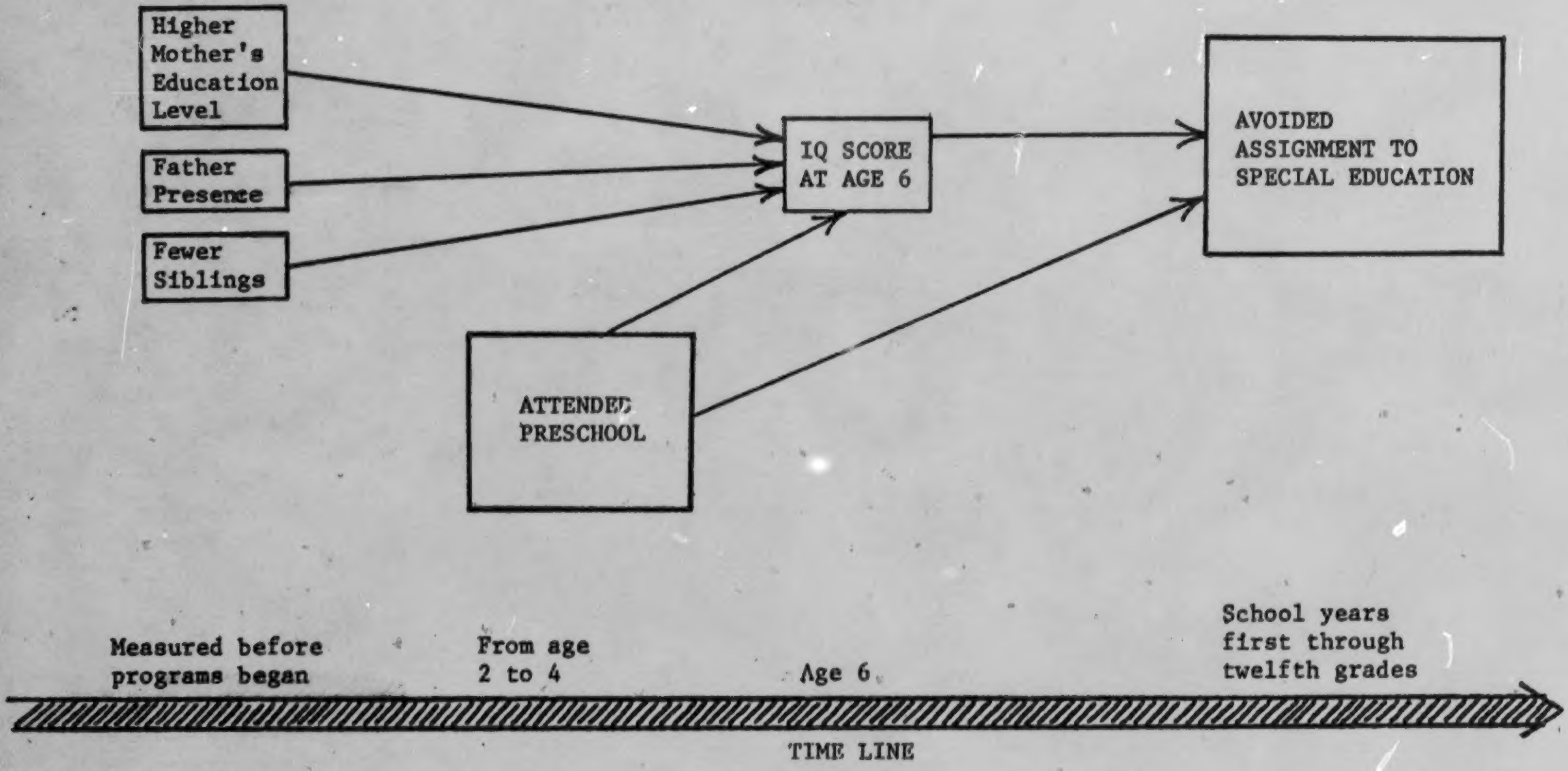
Figure 1 represents a diagram of the relationships between early background measures, preschool attendance, IQ score at age 6,

---

[1] Readers more accustomed to research utilizing cross-sectional data might wonder why we did not make use of the age differences among the 2,000 odd subjects to examine relationships among variables in more detail. This procedure was deemed impractical because age of the children is inextricably confounded with project. That is, Gordon's children were aged 9 and Gray's children aged 19 at follow-up. Furthermore, there were cohort differences, with some children entering preschool at the height of the War on Poverty and others entering as the Nixon Administration began dismantling many programs.

[2] IQ scores at 6 years old were Stanford-Binet in all cases except PPVT scores were used for Levenstein's sample. Palmer's children took the Stanford-Binet at age 5.

Figure 1: Assignment to Special Education Classes: Diagram Showing Network of Variables Suggested by Data, Significant Paths Only


Higher Mother's Education Level
Father Presence
Fewer Siblings
IQ SCORE AT AGE 6
AVOIDED ASSIGNMENT TO SPECIAL EDUCATION
ATTENDED PRESCHOOL
Measured before programs began
From age 2 to 4
Age 6
School years first through twelfth grades
TIME LINE

and later school outcome, in this case assignment to regular vs. special education classes. A time line has been drawn in to indicate the child's age at each measurement period. Each of the links pictured in the diagram represents an hypothesis test reported in the body of this report or in Royce (1979). Let us use this diagram to guide us through a discussion of the impact of preschool on low-income children.

On the far left are the variables measured before children enrolled in preschool. These are our most direct indicators of the early status of the children's background and their intellectual potential. Although not pictured in Figure 1, family background measures were related to the child's early IQ score, consonant with other research. Children from two-parent homes with fewer siblings and with mothers who completed more years of school were more likely to score high on IQ tests administered at age 3 or 4.[1] Limited and controversial though they may be, IQ test scores do provide a measure of cognitive ability and, furthermore, are predictive of later school performance. Thus, the relationship of background variables to early IQ scores indicates that even within a lower-income group, some children started out "ahead" of others.

Many of these children then participated in preschool programs of various kinds. The next time we assessed them as a group was at age 6, on the threshold of the first grade. Again, the measure of cognitive ability was an IQ test score. The reader will note the arrows connecting the background variables with IQ score at 6. These arrows indicate that in one sense the picture is the same as it was before; namely, children from two-parent families with few siblings and with more educated mothers scored higher on this measure of cognitive ability.[1] Notice, however, that the preschool attendance variable also connects with the IQ-at-age-6 variable. In other words, preschool became a new factor in these children's lives. Attending preschool also predicted a higher IQ score at age 6. Home background and preschool attendance were both important influences. If the effects of preschool were partialled out, the background variables still predicted higher IQ scores at age 6. And, vice-versa, if the effects of the background variables were partialled out, preschool attendance still predicted higher IQ scores.

---

[1] The relationship between mother's education and pretest IQ and number of siblings and pretest IQ were reported in Murray (1977). For the relationship between father presence vs. absence and pretest IQ, the correlation matrices were scrutinized. r's varied from -.09 to .24 across eight projects.

We next assessed the group of children in the 1976-77 follow-up study. They ranged from 9 to 19 years old and either had completed their school careers or were enrolled in grades three through twelve (or, in some cases, had dropped out of school). This time the dependent measure of interest was assignment to regular vs. special education classes. What is the relationship between the children's early background, preschool attendance, and their later school careers (i.e., avoiding placement in special education classes)?[1] Now the picture is quite different. Not surprisingly, children's IQ scores at age 6 strongly predicted their school performance. In addition, preschool attendance predicted avoiding placement in special education, even if the effect of preschool on IQ score at age 6 was partialled out. The home background variables have dropped out of the picture, however. There was a relationship between mother's education and child's later school performance, but it disappeared when the effects of IQ at age 6 were partialled out. Furthermore, we have additional information about preschool attendance and family background that is not, for simplicity's sake, drawn into the diagram. We know that preschool helped all types of low-income children avoid placement in special education, regardless of family structure, family size, maternal education, sex of child, ethnic background, or initial IQ score of the child. Therefore, it seems safe to say that by school age, IQ scores at age 6 and preschool attendance importantly affected later school performance, as measured by children's placement in regular vs. special education classrooms.

---

[1] This is true even after the effects of initial IQ scores on IQ scores at age 6 was partialled out. For five data sets with pretest IQ, the pooled p values predicting IQ at 6 were .0001, .024, .025, .061 for mother's education, father present, number of siblings, and sex respectively. For the seven projects, the pooled p values predicting IQ at 6 were .001, .018, .017, .750 for mother's education, father present, number of siblings, and sex respectively.

These are striking findings and worthy of careful consideration. But many questions remain to be answered. The reader will recall, for example, that the effects of preschool attendance on retention in grade were not so large and that preschool attendance did not predict later retention in grade independently of its effect on IQ score at age 6. We have some reason to believe that the retention-in-grade variable is a weak one, but it may be that preschool just does not have the same impact on this variable. If that were so, why should it be the case?

---

[1] At this point we have only analyzed whether children had ever been assigned to special education classes (or retained in grade). We plan further analyses to ascertain when children were so assigned or retained. Preliminary scrutiny of these data lead us to believe that results will not be substantially different.

Furthermore, we have only scratched the surface with our measures; there is a plethora of unmeasured intervening variables in need of investigation in order to clarify the process by which preschool exerted its impact. By partialling out the effect of preschool on IQ score at age 6, we essentially found that preschool affected children in ways that were relevant to school performance but not related to cognitive skills and abilities. Perhaps children's achievement motivation, values, aspirations, or coping styles were influenced. We reported limited evidence that this was so earlier in this report. Perhaps children's classroom behaviors were affected. Individual investigators (e.g., Beller, 1974) have reported that teacher ratings of children with preschool experience differed from those of control children. Children's families may have been influenced by, for example, changing parents' perceptions of their children, affecting the family dynamics, increasing their hopes for the children's future. Again, we reported limited evidence that preschool affected maternal aspirations. But we have virtually no evidence about the influence of the larger social and historical context. How did desegregation and busing enter into this picture? What difference did it make to enroll a child in intervention programs at the height of a societal commitment to social change? What will be the effect of the current disillusion with social legislation and spending? To answer these questions, investigators must continue to design and carry out further longitudinal studies. The Consortium for Longitudinal Studies has provided a baseline from which to operate by demonstrating that preschool intervention programs can make a lasting difference in lives of low-income children.

SUMMARY AND IMPLICATIONS

This report contains the analyses of longitudinal data designed to measure the effects of preschool intervention programs on low-income children. These analyses were conducted by the Consortium for Longitudinal Studies, a group of 14 investigators, 12 of whom had designed and operated intervention programs in the 1960s. The original investigators had hoped to increase low-income children's ability to perform adequately in school. Many also had goals such as influencing children's and families' attitudes toward school, increasing feelings of self-worth, etc. In this respect they were similar to (and some were harbingers of) Head Start. They were dissimilar in that all the programs were also designed as research projects. Now, 10 to 15 years later (depending on the inception of each program) it is possible to evaluate whether and to what extent these preschool intervention programs had any impact on children's lives.

Let us consider first the goal of increasing children's school performance and the (sometimes more implicit) goal of influencing their cognitive ability. The Consortium has three kinds of measures relevant for evaluating the preschool programs' success in terms of these goals: IQ test scores, children's school status during their school careers, and children's achievement test scores.

Children participating in preschool programs and the control children were given IQ tests at least three times: before inception of the program, at age 6, and in a 1976-77 follow-up study when they were 9 to 19 years old (depending on the particular program they attended). The results of comparing treatment and control children were very clear. Before preschool experience, the IQ scores of the two groups did not differ. At age 6, treatment children scored significantly higher than controls on IQ tests. Treatment children maintained this superiority for at least 3 years after the end of the preschool program (cf. Consortium, 1977). By the time of the 1976-77 follow-up, however, there were no significant differences between treatment and control children on WISC-R scores (including the full, verbal, and performance IQ scores and the WISC-R subtest scores) in most projects. The Levenstein and Palmer projects, whose children were less than 13 years old, did find treatment/control differences. We tested the possibility that treatment effects on WISC-R scores were masked because only some kinds of children were affected. There was no evidence, however, that preschool affected WISC-R scores of girls more than boys or that it differentially affected children whose mothers had completed more vs. less education. Nor was there any evidence that preschool helped some children and hurt others (by, for example, making later school classes seem dull in comparison). Thus, we can conclude that preschool programs resulted in short term gains in IQ scores for a period of at least 3 years after the

---

[1] Family structure (father absence vs. presence), family size (number of siblings) and level of education completed by the child's mother.

program ended, but that those gains were not in evidence in those projects whose children were 13 years old or older.

IQ scores serve as an operational definition of intelligence or cognitive ability; but this use of intelligence tests has been the subject of much controversy. Since the original intervention programs were aimed at influencing children's school performance, the Consortium gathered information from the schools about children's school careers. Children who avoided grade failure (retention in grade) and avoided assignment to special education classes were clearly able to meet at least the minimum requirements of their schools. Let us now examine the evidence that preschool affected these basic measures of school performance.

Findings were strong regarding the effects of preschool on these measures of school performance. Children who attended preschool were only about half as likely as control children to be assigned to special education classes. They were also less likely to be retained in grade. These results also held true when we controlled for effects of children's initial IQ score, sex, ethnic background, and three measures of family background (recorded prior to the program).

Furthermore, there was evidence that all these low-income children were helped; that is, benefits were not limited to children with higher initial IQ scores, to boys vs. girls, to white vs. Black children, to children from two-parent vs. one-parent homes, to children with fewer vs. more siblings, or to children whose mothers had more vs. less education.

Some critics have maintained that preschool effects on school performance come about through temporarily raising IQ scores, which thus causes teachers to label these children as brighter. To test this hypothesis, we partialled out the effects of preschool on IQ scores at age 6. Even when controlling for preschool's effect on cognitive ability (as measured by the IQ test), the preschool variable still independently predicted assignment to special education classes (with treatment children less likely to be so assigned).

Generally, results from retention in grade were somewhat weaker than those for assignment to special education. In retrospect, we attribute this to two reasons: (a) many school districts had automatic promotion policies, with a resultant smaller variation among children; (b) a higher proportion of control children were assigned to special education classes; once in such classes they were less likely to return to regular classes and be retained in grade.

Achievement test scores provide another measure of school performance, insofar as they reflect the content children learned in their classrooms. Our analyses of achievement tests are in the early stages, but already it is evident that treatment children scored higher than control children at the fourth grade level. The difference between the groups was significant for math achievement tests; in the median project, treatment children were about a half-grade ahead of the control children.

In summary, it appears that the preschool programs directed by Consortium members had substantial and lasting effects on the school performance of low-income children. Treatment children were more likely to meet the minimum requirements of their schools. They were also more likely to score higher on standardized math achievement tests in the fourth grade.

Once having established that preschool programs affected later school performance, we turned to questions of program characteristics. What kinds of programs had the best records in terms of enabling children to later avoid assignment to special education? Did certain kinds of programs seem to work better for some kinds of children? For example, did girls benefit more than boys from higher adult/child ratios? Ten different program characteristics were examined in this way: age of children at intervention, the length of the intervention (in both years and months per year), the number of hours per year of instruction, the level of parent involvement, the presence or absence of language goals, the existence of preservice training for teachers, the degree of structure in the teaching methods, the location of the program (center vs. home), and staff characteristics (professional vs. non-professional). In these, our most recent analyses, we could find little evidence that some characteristics were more helpful than others. Furthermore, there was no evidence that these characteristics were differentially successful with children who varied by sex, initial IQ score, ethnic background, family structure and size, and maternal educational level. While there was variation among the programs in terms of the ten program characteristics, they were similar in the sense that all were well-run, high quality programs. In brief, then, it appears that all low-income children can benefit from well-run preschool programs when the measure of program effectiveness is assignment to regular vs. special education classrooms.

Finally, we explored the question of how preschool programs exerted their effect. As Gray pointed out (Klaus & Gray, 1974), it is highly unlikely that an "innoculation" of preschool could protect children from school failure over the ensuing 12 years. Furthermore, the fact that preschool attendance predicted assignment to regular classrooms even after controlling for its influence on cognitive ability suggested that non-cognitive aspects of development were affected as well. We considered four areas in detail as possibly showing evidence of persisting preschool influence: maternal aspirations for their children, children's later achievement orientation, children's self-evaluation, and children's sociability. There were strong results for two dependent variables. First, mothers of children who had attended preschool had much higher vocational aspirations for their children than the childrn had for themselves (discrepancy score). This pattern was not true for mothers of control children and their offspring. Secondly, children who had attended preschool were much more likely to mention achievement-related reasons for feeling proud of themselves. This was especially true for girls.

Another exploratory analysis was undertaken to determine if families of treatment and control children differed in their use of Title IV child welfare services. Although the method of using archival records from the various states proved workable, no significant differences emerged in the use of these services.

Implications

The first report of the Consortium for Longitudinal Studies (Consortium, 1977) concluded that preschool intervention programs helped low-income children meet the minimal requirements of their respective schools, either by reducing the rate of placement in special education or by avoiding grade failure (retention in grade). Most of our work over the past year constitutes an attempt to understand the mechanisms, or to limit or qualify, that earlier general conclusion.

Using the children's school outcomes (i.e., assignment to special education classes, retention in grade) as dependent variables, we questioned whether some kinds of children benefitted more from preschool than others.

The answer was no.

We queried whether some program characteristics were more successful than others in reducing later assignments to special education classes.

The answer, using these data, was no.

We tested whether some program characteristics were more successful with certain kinds of children using assignment to regular vs. special education classes as the criterion for success.

Again, the answer was no.

Whether controlling for IQ at age 6 -- when the effects of preschool attendance on IQ scores are at their greatest -- or for

IQ prior to intervention, the experimental-control differences in special education placement remained significant. The same was true when we controlled for background variables.

Thus, this report emerges with the same general conclusion, buttressed by more extensive evidence: all these high-quality preschool programs apparently benefitted their low-income participants by enabling them to meet the minimal requirements of their school systems. In addition, we have added to this conclusion another indication of how the children were helped: children who attended preschools scored higher than control children on math achievement tests in the fourth grade with a suggestive trend toward scoring higher on reading achievement tests as well.

The fact that we found no interactions between children's early family backgrounds, their early intellectual abilities, and their later school outcomes will surprise many scholars. And, of course, policy makers will be disappointed that no one program characteristic emerged as superior in these analyses. We ourselves did not expect these results. To some extent, these negative findings were due to the dependent variables used. Assignment to special education and retention in grade are molar measures. They are also dichotomous: children either did or did not benefit. It may be necessary to use more sensitive outcome measures in order to detect interaction effects. Our future plans include repeating all these analyses using achievement test scores as the outcome measure.

We should also like to point out, however, the strengths of the outcome measures we used. Assignment to special education and retention in grade are meaningful at the level of the individual child; that is, s/he faces the possibility of experiencing either one or both, and these events have important real-life consequences for him or her, as we detailed earlier in this report. The measures are also useful in a social sense. They are comprehensible to laypeople as well as to professionals. Furthermore, it is possible to assign dollar values to special education programs (and, to a lesser extent, to grade failure) so taxpayers and decision-makers can readily see the benefit of avoiding special education placement and retention in grade in dollar and cents terms. In fact, if legislators were to look to the Consortium on Longitudinal Studies for advice on how to spend taxpayers' dollars, we would answer unequivocally: allocate money to preschool intervention programs. Allocate enough funds (and to the appropriate recipients) to ensure that the programs will be well-run. In terms of enabling lower-income children to avoid placement in special education classes and/or retention in grade, it appears to be a sound social investment.

The other outcome variable included here was IQ test scores, measured in 1976-77 when children ranged from 9 to 19 years old. In general, there were no treatment/control differences on full, verbal, or performance WISC-R IQ scores or on the subtests in those projects whose participants were 13 years of age or older. This result should be considered in the context of finding large and reliable treatment/control differences on the other outcome measures. We believe that this result implies that caution should be exercised in using IQ scores as outcome measures for intervention programs, particularly over a long period of time. These results support Zigler's and Trickett's (1978) recent questioning of the utility and meaning of IQ scores in evaluation research.

Finally, we will mention again that some of the analyses included in this report point to preschools' effects on children's achievement-orientation and on parents' aspirations for their children. This, in conjunction with the treatment/control differences on school outcomes and achievement tests, implies that future research might focus on the non-cognitive (or social competence) area, so that we may be able to specify further effects of preschool and to explore the processes by which preschool exerts its impact.

REFERENCES

Abelson, E.D., Zigler, E., & DeBlasi, C.L. Effects of a four-year follow through program on economically disadvantaged children. *Journal of Educational Psychology*, 1974, *66*, 756-771.

Abt Associates, Inc. *National day care study: Preliminary findings and their implications*. Preliminary report, Contract No. HEW-1[illegible]5-74-1100 to the Administration for Children, Youth and Families, Office of Human Development Services, U.S. Department of Health, Education, and Welfare. January, 1978.

Anderson, J.G., & Evans, F.B. Family socialization and educational achievement in two cultures: Mexican-American and Anglo-American. *Sociometry*, 1976, *39*, 209-222.

Beller, E.K. Impact of early education on disadvantaged children. In S. Ryan (Ed.), *A report on longitudinal evaluations of preschool programs*, Vol. 1: 15-48. DHEW Publication No. (OHD) 74-24, 1974.

Bloom, B.S., Davis, A., & Hess, R. *Compensatory education for cultural deprivation*. New York: Holt, 1965.

Boocock, S.S. *An introduction to the sociology of learning*. New York: Houghton Mifflin, 1972.

Bronfenbrenner, U. *A report on longitudinal programs, Vol. 11: Is early intervention effective*? DHEW Publication No. (OHD) 74-25, 1974.

Bronfenbrenner, U. *The experimental ecology of human development*. Cambridge, Mass.: Harvard University Press, in press.

Brown, B. *Found: Long-term gains from early intervention*. Boulder: Colo.: Westview Press, 1978.

Camilli, G., & Hopkins, K.D. Applicability of chi-square to 2x2 contingency tables with small expected cell frequencies. *Psychological Bulletin*, 1978, *85*, 163-167.

Campbell, D.T. Temporal changes in treatment-effect correlation: A quasi-experimental model for institutional records and longitudinal studies. In V. Glass (Ed.), *Proceedings of the 1970 Invitational Conference on Testing Problems*. Princeton, N.J.: Educational Testing Service, 1971.

Campbell, D.T., & Erlebacher, A. How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In E.L. Struening & M. Guttentag (Eds.), *Handbook of evaluation research*, Vol. 1. Beverly Hills: Sage Publications, 1975.

Cicirelli, V., et al. The impact of Head Start: An evaluation of Head Start on children's cognitive and affective development. A report presented to the Office of Economic Opportunity, pursuant to Contract B89-4536, June, 1969. Westinghouse Learning Corporation, Ohio University. Distributed by: Clearinghouse for Federal Scientific and Technical Information, U.S. Department of Commerce, National Bureau of Standards, Institute for Applied Technology. PB 184 328. (Westinghouse Report)

Coleman, et al. Equality of educational opportunity. Washington, D.C.: Government Printing Office, 1966.

The Consortium on Developmental Continuity. The persistence of preschool effects. Final report, Grant No. 18-76-07843 from the Administration for Children, Youth, and Families, Office of Human Development Services, U.S. Department of Health, Education, and Welfare. October, 1977.

Darlington, R.B. Methods, issues, and some illustrative findings in analyzing the data of the Consortium on Developmental Continuity. Paper presented at the Annual Meetings of the American Educational Research Association, Toronto, March 27, 1978a.

Darlington, R.B. Multiple comparisons and related problems. Workshop presented at the Ammual Meetings of the American Psychological Association, Toronto, August, 1978b.

Darlington, R.B. Radicals and squares. Ithaca, N.Y.: Logan Hill Press, 1975.

Datta, L-E. The impact of the Westinghouse/Ohio evaluation on the development of Project Head Start: An examination of the immediate and long-term effects and how they came about. In Abt, C. (Ed.), The evaluation of social programs. Beverly Hills, Calif.: Sage Publications, 1976.

Datta, L-E. What has the impact of Head Start been: Some findings from national evaluations of Head Start. Undated manuscript.

Garber, H., & Heber, F.R. The Milwaukee Project: Indications of the effectiveness of early intervention in preventing metnal retardation. In P. Mittler (Ed.), Research to practice in mental retardation: Care and intervention. Vol. 1. Baltimore: University Park Press, 1977.

Gordon, I.J., Guinagh, B., & Jester, R.E. The Florida parent education infant and toddler programs. In M.C. Day & R.K. Parker (Eds.), The preschool in action (2nd edition). Boston: Allyn & Bacon, 1977.

Gordon, I.J., & Guinagh, B.J. Middle school performance as a function of infant stimulation. Final report to the Administration for Children, Youth, and Families, Department of Health, Education and Welfare, Project No. NIH-HEW-OCD-90-C-908, March 31, 1978.

Gray, S.W. Children from three to ten: The early training project. In S. Ryan (Ed.), *A report on longitudinal evaluations of preschool programs*, Vol. 1: Longitudinal evaluations. Washington, DHEW Publication No. (OHD) 74-24, 1974.

Gray, S.W., Klaus, R.A., & Ramsey, B.K. The early training project participants at the end of high school. Paper presented at the Annual Meetings of the American Educational Research Association, Toronto, March, 1978.

Gray, S.W., & Klaus, R.A. The early training project: A seventh year report. *Child Development*, 1970, *41*, 909-924.

Herzig, M.E., Birch, H.G., Thomas, A., & Mendez, O.A. Class and ethnic differences in the responsiveness of preschool children to cognitive demands. *Monographs of the Society for Research in Child Development*, 1968, 33, serial no. 117.

Herzog, E., Newcomb, C.H., & Cisin, I.H. Double deprivation: The less they have the less they learn. In S. Ryan (Ed.), *A report on longitudinal evaluations of preschool programs*, Vol. 1. Washington, D.C.: DHEW Publication No. (OHD) 74-24, 1974.

Hollingshead, A.B. Two-factor index of social position. Unpublished Manuscript, 1957.

Hubbell, V.R. Differential effects of early childhood intervention programs. Paper presented at the Annual Meetings of the American Orthopsychiatric Association, New York City, April, 1977.

Jencks, C.M., Smith, H., Ackland, M., Bane, C.D., Gintin, H., Heynes, B., & Michelson, S. *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books, 1972.

Jensen, A.R. How much can we boost I.Q. and scholastic achievement? *Harvard Educational Review*, 1969, reprint series no. 2: 1-123.

Jensen, A.R. The race x sex x ability interaction. In R. Canero (Ed.), Intelligence: Genetic and environmental influences. New York: Grune and Stratton, 1971.

Kahl, J.A. Some measurements of achievement orientation. *American Journal of Sociology*, 1965, *70*, 669-681.

Karnes, M.B., Zehrbach, R.R., & Teska, J.A. Conceptualization of the GOAL (game-oriented activities for learning) curriculum. In M.C. Day & R.K. Parker (Eds.), *The preschool in action: Exploring early childhood programs* (2nd Edition). Boston: Allyn & Bacon, 1977.

Kenniston, K., & Carnegie Council on Children. All our children: The American family under pressure. New York: Harcourt Brace Jovanovich, 1977.

Klaus, R.A., & Gray, S.A. The Early Training Project for disadvantaged children: A report after five years. Monographs of the Society for Research in Child Development, 1968, 33 (4), serial no. 120.

Kohn, M. Social competence, symptoms and underachievement in childhood: A longitudinal perspective. New York: Wiley, 1977.

Koretz, D.M. Long term non-cognitive effects of seven infant and preschool intervention programs. Cornell University, University Microfilms, 1978.

Koretz, D.M., Vopava, J., & Darlington, R. Matching early education programs to children: Some methodological limitations. Paper presented at the American Association for the Advancement of Science, Washington, D.C., February 15, 1978.

Kraus, P.E. Yesterday's children: A longitudinal study of children from kindergarten into adult years. New York: Wiley & Sons, 1973.

Ladner, J.A. Tomorrow's Tomorrow: The Black woman. Garden City, N.Y.: Doubleday, 1971.

Levenstein, P. Third grade effects of the mother-child home program. Final report to the Developmental Continuity Consortium, April 30, 1978.

Levenstein, P. The mother-child home program. In M.C. Day & R.C. Parker (Eds.), The preschool in action (2nd Edition). Boston: Allyn and Bacon, 1977.

Lewis, M., & McGurk, H. Evaluation of infant intelligence. Science, 1972, 178, 1174-1177.

McClelland, D.C. Measuring motivation in phantasy: The achievement motive. In D.C. McClelland (Ed.), Studies in motivation. New York: Appleton-Century-Crofts, 1955.

McClelland, D.C. Testing for competence rather than for "intelligence". American Psychologist, 1973, 28, 1-14.

McKay, H., et al. Improving cognitive ability in chronically deprived children. Science, 1978, 200, 270-278.

Madden, J., Levenstein, P., & Levenstein, S. Longitudinal IQ outcomes of mother-child home program. Child Development, 1976, 47, 1015-1025.

Mercer, J. Psychological assessment and the rights of children. In N. Hobbs (Ed.), Issues in the classification of children, Vol. 1. San Francisco: Jossey-Bass, 1975.

Miller, L.B. Highlights of sixth--seventh grade follow-up on achievement and IQ for longitudinal intervention study. Memorandum to Developmental Continuity Consortium members, March 31, 1978.

Milofsky, C.D. Why special education isn't special. Harvard Educational Review, 1974, 45.

Moriarity, A. Coping patterns of preschool children in response to intelligence test demands. General Psychology Monographs, 1961, 64, 3-127.

Mosteller, F., & Bush, R.R. Selected quantitative techniques. In G. Lindzey (Ed.), Handbook of social psychology, Vol. 1. Reading Mass.: Addison-Wesley, 1954.

Mosteller, F., & Tuken, J. Data analysis and regression. Reading Mass.: Addison-Wesley, 1977.

Murray, H.W. Early intervention in the context of family characteristics. Paper presented at the Annual Meetings of the American Orthopsychiatric Association, New York, April, 1977.

Ogbu, J.U. Minority education and caste: The American system in cross-cultural perspective. New York: Harcourt Brace Jovanovich, 1978.

O'Leary, S., & Schneider, M.R. Special class placement for conduct problem children. Exceptional children, 1977, 24-30.

Palmer, F.H. The effects of minimal early intervention on subsequent school performance and I.Q. Final Report, Contract No. 26-77-09617 Education Commission of the States. New York: Palmer Associates, 1977.

Palmer, F.H., Semlear, T., & Fisher, M. One-to-one. In M.J. Begab, H.C. Haywood, & H.L. Garber (Eds.), Prevention of retarded development in psychosocially disadvantaged children. Baltimore: University Park Press, 1978.

Palmer, F.H., & Semlear, T. Early intervention: The Harlem Study. In R. Liebert, R. Poulus, & G. Marmor (Eds.), Developmental Psychology. New York: Prentice-Hall, 1976.

Palmer, F.H., & Siegel, R.J., with Persons, W. & Maltese, M. Harlem study follow-up: Cognitive and social characteristics of the sample at ages 13 to 15. Final Report, Grant No. 90-C1274(01) from the Administration for Children, Youth and Families to Early Intellective Development, Inc., October, 1978.

Palmer, F.H., & Siegel, R.J. Minimal intervention at ages two and three and subsequent intellective changes. In M.C. Day & R.K. Parker (Eds.), The preschool in action: Exploring early childhood programs (2nd Edition). Boston: Allyn & Bacon, 1977.

Palmer, F.H., & Siegel, R.J. Harlem Study follow-up: Cognitive and social characteristics of the sample at ages 13 to 15. Final Report, Grant No. 90-C1274(01) from the Administration for Children, Youth and Families (OHDS; DHEW). October, 1978.

Raizen, S., Bobrow, S.B., Bikson, T.K., Butler, J.A., Heald, K., & Ratteray, J. Design for a national evaluation of social competence. Santa Monica: The Rand Corporation, 1974. Grant No. H-9766-A/H/O, Office of Child Development, U.S. Department of Health, Education, and Welfare. (ERIC Document Reproduction Service No. ED 108 569).

Rosché, M. Early intervention and later use of child welfare services. Unpublished master's thesis, Cornell University, 1978.

Royce, J. The long-term effects of background characteristics and early education programs on later school outcomes. Cornell University, University Microfilms, 1979.

Schulz, D.A. Coming up black: Patterns of ghetto socialization. Englewood Cliffs, N.J.: Prentice-Hall, 1969.

Schweinhart, L.J., & Weikart, D.P. Perry Preschool effects nine years later. What do they mean? Paper presented at the NICHD Conference on Prevention of Retarded Development in Psychosocially Disadvantaged Children, University of Wisconsin, Madison, July, 1978.

Sowell, T. Arthur Jensen and his critics: The great I.Q. controversy. Change, 1973, 5, 33-37.

Spenner, K.I., & Featherman, D.L. Achievement ambitions. Discussion paper DP#449-77. Madison, Wisconsin: Institute for Research on Poverty, 1977. (To appear in Vol. 4 of the Annual Review of Sociology, 1978).

Stack, D. All our kin. New York: Harper Row, 1974.

Vopava, J., & Royce, J. Comparison of the long-term effects of infant and preschool programs on academic performance. Paper presented at the Annual Meetings of the American Educational Research Association, Toronto, March 27, 1978.

Weber, C.U., Foster, P.W., & Weikart, D.P. An economic analysis of the Ypsilanti Perry Preschool project. Monographs of the High/Scope Educational Research Foundation, No. 5, 1978.

Weikart, D.P., Bond, J.T., & McNeil, J.T. The Ypsilanti Perry Preschool Project: Preschool years and longitudinal results. Monographs of the High/Scope Educational Research Foundation, No. 3, 1978.

Weikart, D., Deloria, D., & Lawser, S. Longitudinal results of the Ypsilanti Perry Preschool Project. Monographs of the High/Scope Educational Reserach Foundation, No. 1, 1978.

Ypelaar, A. An analysis of preschool program characteristics and placement in special education: A second look. Unpublished manuscript, 1978.

Zigler, E., & Trickett, P.K. IQ, social competence, and evaluation of early childhood education programs. American Psychologist, 1978, 33, 789-798.

APPENDIX A

Attrition

One of the most serious threats to any longitudinal study is the problem of attrition -- the fact that, over time, some of the subjects drop out of a study (due to moving, death, or a myriad of other reasons).

The analysis of whether attrition has caused biases in the final sample will be directed towards answering four specific questions:

1. Are there different rates of attrition for treatment and control groups?

2. Do the final samples differ on some important characteristic from the dropouts?

3. Are there any instances of differential attrition? That is, are different kinds of children selectively retrieved in experimental groups than in control groups?

4. Do final program samples and final control samples differ on some important characteristics?

The first three of these questions are questions on attrition in the strict sense of the term, That is, they answer the question of the extent to which final sample treatment and control groups represent the original sample groups. The usefulness of such information in large measure hinges on the demonstrated -- or assumed -- initial equivalence of the treatment and control groups. If they were equivalent initially, attrition analyses provide a method of assessing the extent to which the final treatment and control groups remain equivalent. This essential underlying issue of the equivalence of the final treatment and control groups is addressed directly by our question 4. Accordingly, while question 4 is not, strictly speaking, a question of attrition, it complements the attrition analysis per se and provides perhaps the strongest test of the degree of sampling bias in the Consortium's final follow-up samples.

The Consortium follow-up involved the collection of four types of information, or instruments: individual intelligence tests, interviews of children and mothers, and school records. For purposes of this report, attrition is defined as the failure to report information for a particular child on a particular instrument. This definition is applied because the different methods for collecting information occasionally resulted in considerable divergence among the samples receiving different instruments. For example, collection of Youth Interview and WISC-R data required actual contact with the child, while collection of the School Record Form data required, instead, actual contact with the child's school records -- a feat which could be either more or less difficult than actual contact with the child, depending on the cooperation and organization of the school district.

Virtually all of the attrition was due to simple inability to locate subjects or other sources of data, such as parents or school records. Less than 3% of our subjects or parents refused to participate.

Attrition in the Current Follow-up

There are five measures of attrition in the current follow-up:

- whether a Parent Interview was reported;
- whether a Youth Interview was reported;
- whether a School Record Form was reported;
- whether a WISC-R was reported; and
- whether any of the above was reported (referred to hereafter as "general attrition").

All four of the questions above have been addressed with regard to each of these five measures of attrition. For each of the five measures, questions 2, 3, and 4 have been answered with respect to three potentially important covariates: pretest IQ, mother's education at the time of the child's entry into preschool, and household SES (Hollingshead Two-Factor Index of Social Position) at the time of the child's entry into preschool. For the sake of clarity, the discussion here will focus primarily on the measures of "general attrition."

The results here are reported on nine projects: Beller, Deutsch, Gordon, Gray, Karnes, Levenstein, Miller, Palmer, and Weikart.[1] The percentages of children found (on any instrument) are given in Table A-1. Our definition of original data is also found in this table. We have excluded certain groups from analyses in this report

---

[1] Woolman's project has no background demographic data of the sort used in the present analysis, and Zigler's project has very little such data that fit the Consortium format.

(unless otherwise noted), and thus have decided to exclude them from our attrition analysis. Gray's group 4 (the "distal control group") differed from her other (randomly assigned) groups in several respects and hence has been excluded. Recently completed analysis of retrospective demographic data from Beller's project indicated that group 3 (first grade without kindergarten or preschool) was dissimilar to the other groups. We have excluded Miller's groups seven and eight since we have no original demographic data for them. And for Karnes' project, we have reported data only for her groups one through seven, since these are the planned curriculum variation groups which we have used for our program variable analyses.

The groups used in the school record form analyses are different from those used in the other attrition analyses for two projects, Gordon and Levenstein. In order to compare children within the same school system, we have excluded from Gordon's project children living outside Alachua County. In Levenstein's project we have excluded group 19 since this is a control group added on at first grade. These exclusions correspond to the data which were actually used in the special education and retention analysis.

The analyses reported here and some further analyses showed the attrition problems for the Deutsch project to be far more severe than those for other projects. Not only was the attrition rate far higher, but the Deutsch team recovered primarily the best treatment-group subjects and the worst control-group subjects. These problems are described in detail in the section of this report describing individual projects. As a result, we were forced to omit the Deutsch data from all of our most important analyses. Therefore the rest of the text in this section will consider primarily the eight projects remaining after the Deutsch project is deleted, although the tables in this section include the Deutsch project.

Table A-1 shows that with the Deutsch project excluded, recovery rates ranged from 35.8% to 100%, with a median of 74.5%. Given the long time intervals involved -- up to 14 years after the completion of preschool -- the average seems reasonably good. Tables A-2 - A-5 show the comparable figures for specific dependent variables. The picture is much the same as in Table A-1.

Tables A-1 - A-5 also show the results of significance tests testing whether recovery rates differed for treatment and control groups. Altogether 32 such tests were performed (excluding Deutsch), so we would expect to find one or two significant differences just by chance, but no differences significant at the .05 level were found. In fact, in these 32 tests no p values were observed below .15, so none of the differences even approached significance. This provides a very clear answer to question 1 in our opening list -- were different rates of attrition found for treatment and control groups? The answer is no.

With respect to general attrition (Table A-1), all projects found roughly the same percentage of control children as they did of program children (i.e., there appears to be no indication of different rates of attrition between program and control groups). Thus, the answer to the first question is negative. However, the percentage found does vary widely among projects. The median percentage found is 73.8. Tables A-2 - A-5 present similar analyses for the individual instruments.

Questions 2, 3, and 4 are answered in Tables A-6 - A-20. These 15 tables result from crossing the five measures of attrition (Parent Interview, Youth Interview, WISC, School Record, and "general attrition" [any instrument obtained]) with the three covariates mentioned above: pretest IQ (usually, Stanford-Binet), Hollingshead Index of Social Position, and mother's education (in grades completed). (As mother's education enters into the calculation of the ISP in father absent homes, this measure overlaps considerably with mother's education.)

Column 1 in each of these tables reports the significance levels of tests which tested whether the recovered and unrecovered groups differed on any of the background measures (IQ, mother's education, and SES, all measured before the beginning of preschool) -- Question 2 in our list. Altogether Column 1 in Tables A-6 - A-20

---

[1] The only significant differences found were on Deutsch parent interviews and school record forms.

reports 107 hypothesis tests of this sort. Of these 107 p values, we would expect 10 or 11 to be below .10 by chance, and 10 were. Thus the number of significant or nearly significant results is almost exactly what would be expected by chance. The smallest one of the 107 p values was .015. Among 107 independent p values, the probability is .80 that at least one p value would be as small as .015. Thus, by this measure also, the set of p values testing Question 2 is essentially what one would one would expect by chance.

Question 3 asked whether different kinds of subjects were recovered in the treatment-group and control-group samples. For instance, if the brighter treatment-group children and the less bright control-group children were recovered, then in the recovered sample the difference between mean age 3 IQ's of treatment and control children would be larger than in the original sample. By the same token, this difference would be larger for the recovered children than for the lost children. Thus Question 3 concerns a difference between differences, or an interaction in an ANOVA table. If we form a 2 x 2 table of treatment vs. control and lost vs. recovered subjects, and enter mean age 3 IQ scores in the four cells, then an ordinary ANOVA test for interaction will test whether different kinds of children were recovered in the treatment group than in the control group.

Where possible, the interaction analysis just described was performed 15 times for each project. It was performed separately for each of three background variables: age 3 IQ, mother's education, and Hollingshead SES. For each of these three variables, it was performed five times -- once for each of the five kinds of attrition mentioned earlier. Because of missing data in various categories, the total number of interaction tests performed was 76. The p values for these tests are shown in column 2 of Tables A-6 - A-20. Of the 76 p values, 10 were below .10, while seven or eight such values would be expected by chance. The smallest of the 76 p values was .032, which again is well within what one might expect by chance.

Although not shown in this report, similar interaction tests were also performed on three other background variables -- father presence or absence at the beginning of preschool, number of siblings at that time, and Stanford-Binet IQ score upon completion of the project. The results of these analyses are essentially the same as those reported in the last paragraph, with one exception: in the Deutsch project there was a significant ($p = .001$) tendency for school records to be recovered for the treatment-group children measured as brightest at the end of preschool, and for the least bright control children. Thus using the recovered Deutsch data would seriously exaggerate the effectiveness of the Deutsch preschool program. Therefore the Deutsch data was omitted from the summaries of most of our analyses. This is discussed in more detail in this report's section on project descriptions.

Question 4 does not concern attrition per se, but it does concern the types of bias which attrition might produce, so it is included here. In the recovered samples, were there significant differences between treatment and control groups on any of the background variables -- age 3 IQ, mother's education, or Hollingshead SES? Again, when data were available, we tested this question in 15 different ways for each project. The p values for these tests appear in column 3 of Tables A-6 - A-20. The 93 tests thus performed yielded eight p values below .10, while 9 or 10 would be expected by chance. However, seven of eight p values were .011 or below, which is low enough to require some discussion.

The Levenstein sample showed p values of .085, .011 and .011, all indicating a difference between recovered treatment and control groups on Hollingshead SES. However, in our most important analysis, the Levenstein project was classified as nonexperimental, thus limiting the potential biases introduced by this problem. Also, there was no hint of such bias in mother's education, which of course correlates with SES. When it is recalled that a p value of .011 could well occur by chance among 93 tests, we are inclined to view this as a rather mild source of bias.

The remaining five small p values in column 3 of Tables A-6 - A-20 all show highly significant differences on age 3 IQ between recovered treatment and control groups in the Palmer project. On "general attrition" (failure to recover any follow-up data concerning a child), the difference is significant beyond the .0001 level, and on the four more specific kinds of attrition the differences are also highly significant. On general attrition, the difference is a 9-point IQ difference favoring the program group. Since Palmer used an essentially random procedure of assigning children to program and control groups, it is difficult to understand how such a large difference could have occurred. Two things lead us to believe that at least part of this difference is spurious. First of all, the control children were tested at an average age of 2 years and 9 months, while the average age of the program children was 3 years. The norming of the Stanford-Binet is dubious at young ages, especially for lower SES children. Secondly, there are no differences for these same children on demographic variables. For example, of the children who were given a pretest IQ, the program children's mothers have completed an average of 11.39 years of education, while the control children's mothers have completed 11.38 years of education. Similarly, the average Hollingshead score for the program children is 58.83; for controls, 57.28. (A higher score indicates a lower socioeconomic status.) These figures lead us to believe that Palmer's program and control groups are more alike than the pretest IQ data indicate. In this report, whenever possible, we have controlled for pretest IQ to help alleviate this initial difference.

Readers who are familiar with the Consortium work may notice some discrepancies between these attrition results and those of our previous report (Consortium, 1977). Any discrepancies are due to two factors. First of all, after our 1977 report we received data on new cases. Secondly, in some cases, we defined the groups to be considered in a different way than before. In all such instances, we have defined final sample program and control groups to produce the most equivalent groups possible.

Conclusions

Some of our central staff members joined this project with serious suspicions concerning attrition and its potential for biasing our major analyses. As a result, our analysis of attrition has been far more extensive than those we have seen in most other longitudinal studies. We related five types of attrition to three background demographic characteristics, examining four different questions related to attrition. We reported altogether over 300 significance tests, any one of which was potentially capable of finding a significant difference between treatment and control groups in some respect. Our total attrition analysis has actually been approximately twice as extensive as that reported here. We have reported here results concerning three background variables -- mother's education, pretest IQ, and SES. We have also largely completed comparable analyses for two other background variables -- family size and father presence or absence -- and have done some attrition-related analyses on IQ at age 6. With minor exceptions discussed in this section, it turned out that it would be hard to imagine a set of results more consistent with the hypothesis that attrition was essentially random, introducing no noticeable biases into our other analyses. The minor exceptions have been considered in our major analyses.

Table A-1

Final Sample As A Percent of Original Sample By Project

Project (Groups)	Final Controls as % of Originals Controls	Final Program as % of Original Program	Total Final Sample as % of Original Total Sample	X^2*	Significance (two-tailed)	Definition of Original Sample
Beller (1-2)	66.0**(35)‡ (53)#	66.1(39) (59)	66.1(74) (112)	.0001	.9920	Groups 1 and 2 = 112 cases
Deutsch (1,2)	16.1(31) (192)	19.2(60) (312)	18.1(91) (504)	.7645	.3819	First 4 waves, Groups 1 and 2, only = 504 cases
Gordon	35.8(24) (67)	34.3(83) (242)	34.6(107) (309)	.0538	.8166	All 309 cases with test scores sent to Cornell
Gray (1-3)	90.5(19) (21)	81.8(36) (44)	84.6(55) (65)	.8186	.2656	Groups 1 thru 3 = 65 cases
Karnes (1-7)		86.3(88) (102)	86.3(88) (102)	--	--	Groups 1 thru 7 = 102 cases
Levenstein	77.9(53) (68)	74.2(135) (182)	75.2(188) (250)	.3763	.5396	All 250 cases sent to Cornell (first 5 waves)
Miller (1-5)	52.9(18) (34)	50.9(109) (214)	51.2(127) (248)	.0473	.8278	Groups 1 thru 5 = 248 cases
Palmer	71.6(48) (67)	74.4(180) (242)	73.8(228) (309)	.2034	.6520	All 309 cases sent to Cornell
Weikart	100(65) (65)	100(58) (58)	100(123) (123)	--	--	All 123 Perry cases sent to Cornell
Total##	69.9(262) (375)	63.7(728) (1143)	65.2(990) (1518)			

* Without Yates correction. See Camilli and Hopkins (1978).

** Percentage of original.

‡ Follow-up sample size.

# Original sample size.

## Deutsch sample is excluded from total.

Table A-2

Final Sample As A Percent of Original Sample for Parent Interview by Project

Project (Groups)	Final Controls as a % of Original Controls	Final Program as a % of Original Program	Total Final Sample as a % of Original Total Sample	X^{2}[*]	Significance (two-tailed)
Beller (1,2)	52.8[**](28)[‡] (53)[#]	64.4(38) (59)	58.9(66) (112)	1.5460	.2137
Deutsch (1,2)	9.9(19) (192)	17.0(53) (312)	14.3(72) (504)	4.8812	.0272
Gordon	35.8(24) (67)	34.3(83) (242)	34.6(107) (309)	.0538	.8166
Gray (1-3)	81.0(17) (21)	81.8(36) (44)	81.5(53) (65)	.0071	.9328
Karnes (1-7)		85.3(87) (102)	85.3(87) (102)	--	--
Levenstein	33.8(23) (68)	41.2(75) (182)	39.2(98) (250)	1.1329	.2871
Miller (1-5)	52.9(18) (34)	50.9(109) (214)	51.2(127) (248)	.0473	.8278
Palmer	46.3(31) (67)	46.3(112) (242)	46.3(143) (309)	<.0001	.9986
Weikart	86.2(56) (65)	86.2(50) (58)	86.2(106) (123)	.0001	.9932

* Without Yates' correction. See Camilli and Hopkins (1978).

** Percentage of original.

± Follow-up sample size.

# Original sample size.

Table A-3

Final Sample As A Percent of Original for Youth Interview, By Project

Project (Groups)	Final Controls as a % of Original Controls	Final Program as a % of Original Program	Total Final Sample as a % of Original Total Sample	X^2*	Significance (two-tailed)
Beller (1,2)	58.5**(31)‡ (53)#	59.3(35) (59)	58.9(66) (112)	.0080	.9288
Deutsch (1,2)	14.1(27) (192)	17.6(55) (312)	16.3(82) (504)	1.1093	.2922
Gordon	28.4(19) (67)	30.2(73) (242)	29.8(92) (309)	.0820	.7746
Gray (1-3)	81.0(17) (21)	75.0(33) (44)	76.9(50) (65)	.2837	.5943
Karnes (1-7)	--	79.4(81) (102)	79.4(81) (102)	--	--
Levenstein	36.8(25) (68)	27.5(50) (182)	30.0(75) (250)	2.0354	.1537
Miller (1-5)	52.9(18) (34)	50.9(109) (214)	51.2(127) (248)	.0473	.8278
Palmer	47.8(32) (67)	46.3(112) (242)	46.6(144) (309)	.0462	.8298
Weikart	87.7(57) (65)	81.0(47) (58)	84.6(104) (123)	1.0402	.3078

---

* Without Yates correction. See Camilli and Hopkins (1978).

** Percentage of original.

‡ Follow-up sample size.

# Original sample size.

Table A-4

Final Sample As A Percent of Original for School Record Form, By Project

Project (Groups)	Final Controls as a % of Original Controls	Final Program as a % of Original Program	Total Final Sample as a % of Original Total Sample	X^{2*}	Significance (two-tailed)
Beller (1,2)	62.3**(33)‡ (53)#	62.7(37) (59)	62.5(70) (112)	.0024	.9609
Deutsch (1,2)	6.3(12) (192)	12.5(39) (312)	10.1(51) (504)	5.1048	.0239
Gordon##	23.2(13) (56)	29.8(68) (228)	28.5(81) (284)	.9636	.3263
Gray (1-3)	90.5(19) (21)	81.8(36) (44)	84.6(55) (65)	.8186	.3656
Karnes (1-7)	--	76.5(78) (102)	76.5(78) (102)	--	--
Levenstein##	56.1(23) (41)	57.1(104) (182)	57.0(127) (223)	.0149	.9028
Miller (1-5)	52.9(18) (34)	50.0(107) (214)	50.4(125) (248)	.1015	.7500
Palmer	71.6(48) (67)	72.7(176) (242)	72.5(224) (309)	.0310	.8602
Weikart	100.0(65) (65)	100.0(58) (58)	100.0(123) (123)	--	--

* Without Yates correction. See Camilli and Hopkins (1978).

** Percentage of original.

‡ Follow-up sample size.

# Original sample size.

## Gordon: Alachua County School District only; Levenstein, excluding Group 19.

Table A-5

Final Sample As A Percent of Original For WISC-R, By Project

Project (Groups)	Final Controls as a % of Original Controls	Final Program as a % of Original Program	Total Final Sample as a % of Original Total Sample	X^{2*}	Significance (two-tailed)
Beller## (1,2)	---	---	---	--	---
Deutsch (1,2)	14.6(28) (192)	17.3(54) (312)	16.3(82) (504)	.6476	.4210
Gordon	28.4**(19)‡ (67)#	29.3(71) (242)	29.1(90) (309)	.0244	.8759
Gray (1-3)	85.7(18) (21)	77.3(34) (44)	80.0(52) (65)	.6331	.4262
Karnes (1-7)	--	77.5(79) (102)	77.5(79) (102)	--	--
Levenstein	36.8(25) (68)	27.5(50) (182)	30.0(75) (250)	2.0354	.1537
Miller (1-5)	52.9(18) (34)	50.9(109) (214)	51.2(127) (248)	.0473	.8278
Palmer	41.8(28) (67)	43.0(104) (242)	42.7(132) (309)	.0301	.8623
Weikart	86.2(56) (65)	93.1(54) (58)	89.4(110) (123)	1.5661	.2108

* Without Yates correction. See Camilli and Hopkins (1978).

** Percentage of original.

‡ Follow-up sample size.

# Original sample size.

## Beller is still collecting WISC data. Deutsch scores are WAIS only.

Table A-6

General Attrition-Comparison of Mother's Education

Project (Groups)	Significance Levels			Mean Grades Completed				
				Control		Program		
	Attrition Main Effect	Interaction	Final Samp. P-C Compar.	Dropout	Final Sample	Dropout	Final Sample	Overall Mean
Beller (1,2)	No Data for Dropouts		.309	--	10.69 (26)	--	11.14 (35)	10.95 (61)
Deutsch (1,2)	.773	.505	.700[##]	9.97 (68)	10.35 (17)	10.27 (41)	10.13 (54)	10.12 (180)
Gordon	.161	--	.501	--	9.73 (15)	10.43 (75)	10.03 (67)	10.19 (157)
Gray (1-3)	A[#]	A	.672[##]	9.00 (2)	9.11 (19)	9.67 (6)	8.79 (34)	8.98 (61)
Karnes (1-7)	.277	--	--	--	--	10.54 (13)	10.07 (84)	10.13 (97)
Levenstein	A	.174	.349	10.93 (15)	10.28 (53)	10.33 (46)	10.61 (134)	10.51 (248)
Miller (1-5)	A	A	.331	10.67 (15)	10.29 (17)	10.89 (95)	10.87 (100)	10.82 (227)
Palmer	.251	A	.556[##]	11.42 (19)	11.27 (48)	11.48 (54)	11.13 (178)	11.23 (299)
Weikart	No Dropouts on School Record		.841	--	9.38 (65)	--	9.47 (58)	9.42 (123)

[#] "A" implies the F for this test was less than 1.

[##] Control group has higher mother's educational level.

Table A-7

General Attrition - Comparison of Hollingshead ISP Scores*

Project (Groups)	Significance Levels			Mean ISP Scores				
				Control		Program		
	Attrition Main Effect	Interaction	Final Samp. P-C Compar.	Dropout	Final Sample	Dropout	Final Sample	Overall Mean
Beller (1,2)	No Data for Dropouts		.795 #	--	56.95 (19)	--	57.95 (20)	57.46 (39)
Deutsch (1,2)	.357	.474	.519	65.90 (68)	66.11 (9)	67.02 (41)	64.91 (44)	65.93 (162)
Gordon	.297	--	--	--	--	63.14 (71)	65.10 (40)	63.81 (111)
Gray (1-3)	.185	.126	.615 #	73.00 (2)	69.84 (19)	63.60 (5)	70.52 (31)	69.77 (57)
Karnes (1-7)	.969	--	--	--	--	65.38 (13)	65.28 (80)	65.30 (93)
Levenstein	A ##	.150	.171	64.27 (15)	65.98 (53)	66.37 (46)	64.11 (133)	64.94 (247)
Miller (1-5)	No Data for Controls		--	--	--	62.76 (76)	63.66 (96)	63.26 (172)
Palmer	.062	A	.597 #	55.05 (19)	58.04 (48)	55.98 (53)	59.20 (175)	58.17 (295)
Weikart	No Dropouts on School Record		.376	--	69.09 (64)	--	67.81 (58)	67.92 (123)

---

* The Hollingshead Index of Social Position (ISP) ranges from 11 to 77, with "11" representing the highest social class and "77" representing the lowest social class.

# Control group has lower ISP, i.e. higher SES level.

## "A" implies the F for this test was less than 1.

Table A-8

General Attrition - Comparison of Pretest Stanford-Binet IQ Scores

Project (Groups)	Significance Levels			Mean IQ Scores				
				Control		Program		
	Attrition Main Effect	Interaction	Final Samp. P-C Compar.	Dropout	Final Sample	Dropout	Final Sample	Overall Mean
Beller (1,2)	A#	.071	.493##	87.00 (18)	93.34 (35)	94.16 (19)	91.16 (38)	91.69 (110)
Deutsch (1,2)	.628	.503	.406	91.74 (109)	89.20 (20)	92.32 (222)	92.17 (52)	91.99 (403)
Gordon	No Pretest IQ	--	--	--	--	--	--	--
Gray (1-3)	A	A	.885	78.50 (2)	88.67 (15)	89.88 (8)	89.33 (36)	88.89 (61)
Karnes (1-7)	.837	--	--	--	--	95.21 (14)	94.67 (88)	94.74 (102)
Levenstein*	.197	.260	.728	87.20 (5)	83.59 (22)	81.41 (44)	84.53 (132)	83.82 (203)
Miller (1-5)	No Pretest IQ	--	--	--	--	--	--	--
Palmer	.083	.079	<.0001	86.50 (18)	84.84 (45)	89.86 (29)	95.87 (91)	91.28 (183)
Weikart	No Dropouts on School Record		.373	--	78.54 (65)	--	79.57 (58)	79.03 (123)

"A" implies the F for this test was less than 1.

Control group has higher pretest IQ score.

* PPVT pretest IQ.

Table A-9

Attrition on the Parent Interview - Comparison of Mother's Education

Project (Groups)	Significance Levels: Attrition Main Effect	Significance Levels: Interaction	Significance Levels: Final Samp. P-C Compar.	Mean Grades Completed: Control: Dropout	Mean Grades Completed: Control: Final Sample	Mean Grades Completed: Program: Dropout	Mean Grades Completed: Program: Final Sample	Overall Mean
Beller (1,2)	No Data for Dropouts		.309	---	10.69 (26)	---	11.14 (35)	10.95 (61)
Deutsch (1,2)	.626	.721	1.000	10.05 (75)	10.00 (10)	10.38 (48)	10.00 (47)	10.12 (180)
Gordon	.107	---	.701	---	9.73 (15)	10.43 (75)	10.01 (67)	10.18 (157)
Gray (1-3)	A#	.106	.277##	7.50 (4)	9.47 (17)	9.67 (6)	8.79 (34)	8.98 (61)
Karnes (1-7)	.292	--	--	--	--	10.50 (14)	10.07 (83)	10.13 (97)
Levenstein	.287	A	.583	10.58 (45)	10.13 (23)	10.64 (105)	10.40 (75)	10.51 (248)
Miller (1-5)	A	A	.171	10.67 (15)	10.29 (17)	10.89 (95)	10.87 (100)	10.82 (227)
Palmer	A	A	.696##	11.33 (36)	11.29 (31)	11.26 (121)	11.16 (111)	11.23 (299)
Weikart	A	.058	.370	9.78 (9)	9.32 (56)	8.00 (8)	9.70 (50)	9.42 (123)

# "A" implies the F for this test was less than 1.
## Control group has higher level of mother's education.

Table A-10

Attrition on the Parent Interview - Comparison of Hollingshead ISP Scores*

Project (Groups)	Significance Levels			Mean ISP Scores				
				Control		Program		
	Attrition Main Effect	Interaction	Final Samp. P-C Compar.	Dropout	Final Sample	Dropout	Final Sample	Overall Mean
Beller (1,2)	No Data for Dropouts		.796#	--	56.95 (19)	--	57.95 (20)	57.46 (39)
Deutsch (1,2)	.281	.369	.202	65.85 (73)	67.25 (4)	67.11 (45)	64.60 (40)	65.93 (162)
Gordon	.297	No Data for Controls	--	--	--	63.14 (71)	65.10 (40)	--
Gray (1-3)	A ##	.033	.437#	74.00 (4)	69.24 (17)	63.60 (5)	70.52 (31)	69.77 (57)
Karnes (1-7)	.698	--	--	--	--	64.36 (14)	65.45 (79)	65.29 (93)
Levenstein	.177	A	.085	64.53 (45)	67.70 (23)	64.31 (104)	65.23 (75)	64.94 (247)
Miller (1-5)	.514	No Data for Controls	--	--	--	62.76 (76)	63.66 (96)	63.26 (172)
Palmer	A	A	.441#	57.47 (36)	56.87 (31)	57.76 (120)	59.22 (108)	58.17 (295)
Weikart	.306	A	.482	68.00 (9)	69.27 (55)	65.13 (8)	68.24 (50)	68.48 (122)

* The Hollingshead Index of Social Position (ISP) ranges from 11 to 77, with "11" representing the highest social class and "77" representing the lowest social class.

# Control group has lower ISP, i.e. higher SES level.

## "A" implies the F for this test was less than 1.

Table A-11

Attrition on the Parent Interview - Comparison of Pretest Stanford-Binet Scores

	Significance Levels			Mean IQ Scores				
				Control		Program		
Project (Groups)	Attrition Main Effect	Interaction	Final Samp. P-C Compar.	Dropout	Final Sample	Dropout	Final Sample	Overall Mean
Beller (1,2)	A#	A	.509##	88.64 (25)	93.46 (28)	94.25 (20)	91.03 (37)	91.69 (110)
Deutsch (1,2)	.652	.169	.181	91.88 (116)	86.62 (13)	92.22 (229)	92.67 (45)	91.99 (403)
Gordon	No Pretest IQ.	--	--	--	--	--	--	--
Gray (1-3)	A	A	.982	81.75 (4)	89.23 (13)	89.88 (8)	89.33 (36)	88.89 (61)
Karnes (1,7)	.947	--	--	--	--	94.60 (15)	94.77 (87)	94.75 (102)
Levenstein*	.026	No Data on Follow-up Controls	--	84.26 (27)	--	82.00 (101)	86.11 (75)	83.82 (203)
Miller (1-5)	No Pretest IQ		--	---	--	--	--	--
Palmer	A	A	.000	85.85 (33)	84.73 (30)	93.13 (64)	95.89 (56)	91.28 (183)
Weikart	A	A	.582	78.00 (9)	78.63 (56)	81.00 (8)	79.34 (50)	79.03 (123)

"A" implies the *F* for this test was less than 1.

## Control group has higher pretest IQ score.

* PPVT pretest IQ.

Table A-12

Attrition on the Youth Interview - Comparison of Mother's Education

Project (Groups)	Significance Levels			Mean Grades Completed				
				Control		Program		
	Attrition Main Effect	Interaction	Final Samp. P-C Compar.	Dropout	Final Sample	Dropout	Final Sample	Overall Mean
Beller (1,2)	No Data for Dropouts		.436	--	10.92 (25)	--	11.23 (31)	11.09 (56)
Deutsch (1,2)	.889	.330	.497##	9.96 (70)	10.47 (15)	10.33 (45)	10.06 (50)	10.12 (180)
Gordon	.142	No Data for Control Dropouts	.768	--	9.73 (11)	10.40 (83)	9.92 (60)	10.17 (154)
Gray (1-3)	.015	.209	.289##	6.50 (4)	9.71 (17)	8.00 (9)	9.19 (31)	8.98 (61)
Karnes (1-7)	.318	--	--	--	--	10.40 (20)	10.06 (77)	10.13 (97)
Levenstein	.309	A#	.614	10.40 (43)	10.12 (25)	10.59 (130)	10.40 (50)	10.51 (248)
Miller (1-5)	A	A	.171	10.67 (15)	10.29 (17)	10.89 (95)	10.87 (100)	10.82 (227)
Palmer	A	A	.702##	11.37 (35)	11.25 (32)	11.29 (121)	11.13 (111)	11.23 (299)
Weikart	.303	.037	.272	10.00 (8)	9.30 (57)	8.18 (11)	9.77 (47)	9.42 (123)

"A" implies the F for this test was less than 1.

Control group has higher level of mother's education.

Table A-13

Attrition on the Youth Interview - Comparison of Hollingshead ISP Scores[*]

Project (Groups)	Significance Levels			Mean ISP Scores				
				Control		Program		
	Attrition Main Effect	Interaction	Final Samp. P-C Compar.	Dropout	Final Sample	Dropout	Final Sample	Overall Mean
Beller (1,2)	No Data for Dropouts		.573[#]	--	56.95 (19)	--	59.11 (18)	58.00 (37)
Deutsch (1,2)	.424	.645	.706	65.93 (70)	65.86 (7)	66.75 (44)	65.05 (41)	65.93 (162)
Gordon	A[##]	No Data for Controls	--	--	--	63.37 (78)	64.97 (33)	63.84 (111)
Gray (1-3)	A	.205	.661[#]	74.00 (4)	69.24 (17)	68.13 (8)	69.96 (28)	69.77 (57)
Karnes (1-7)	.857	--	--	--	--	65.60 (20)	65.21 (73)	65.29 (93)
Levenstein	A	.036	.011	64.33 (43)	67.80 (25)	65.16 (129)	63.50 (50)	64.94 (247)
Miller (1-5)	.514	No Data for Controls	--	--	--	62.76 (76)	63.66 (96)	63.26 (172)
Palmer	A	A	.497[#]	57.46 (35)	56.91 (32)	58.02 (120)	58.93 (108)	58.17 (295)
Weikart	.016	A	.672	66.50 (8)	69.46 (56)	63.09 (11)	68.91 (47)	68.48 (122)

---

[*] The Hollingshead Index of Social Position (ISP) ranges from 11 to 77, with "11" representing the highest social class and "77" representing the lowest social class.

[#] Control group has lower ISP, i.e. higher SES level.

[##] "A" implies the F for this test was less than 1.

Table A-14

Attrition on the Youth Interview - Comparison of Pretest Stanford-Binet IQ Scores

	Significance Levels			Mean IQ Scores				
				Control		Program		
Project (Groups)	Attrition Main Effect	Interaction	Final Samp. P-C Compar.	Dropout	Final Sample	Dropout	Final Sample	Overall Mean
Beller (1,2)	.219	.082	.459 ##	86.82 (22)	94.29 (31)	92.83 (23)	91.71 (34)	91.69 (110)
Deutsch (1,2)	.830	.850	.925	91.24 (112)	92.06 (17)	92.27 (226)	92.38 (48)	91.99 (403)
Gordon	No Pretest IQ	--	--	--	--	--	--	--
Gray (1-3)	A #	A	.908	79.67 (3)	89.14 (14)	88.73 (11)	89.67 (33)	88.89 (61)
Karnes (1-7)	.616	--	--	--	--	93.76 (21)	95.00 (81)	94.74 (102)
Levenstein*	.044	No Data on Follow-up Controls	--	84.26 (27)	--	82.62 (126)	86.60 (50)	83.82 (203)
Miller (1-5)	No Pretest IQ	--	--	--	--	--	--	--
Palmer	A	A	.000	85.59 (32)	85.03 (31)	93.24 (66)	95.85 (54)	91.28 (183)
Weikart	A	A	.465	79.75 (8)	78.36 (57)	80.55 (11)	79.34 (47)	79.02 (123)

"A" implies the F for this test was less than 1.
Control group has higher pretest IQ score.
* PPVT pretest IQ.

Table A-15

Attrition on the School Record - Comparison of Mother's Education

Project (Groups)	Significance Levels			Mean Grades Completed				
				Control		Program		
	Attrition Main Effect	Interaction	Final Samp. P-C Compar.	Dropout	Final Sample	Dropout	Final Sample	Overall Mean
Beller (1,2)	No Data on Dropouts		.313	--	10.72 (25)	--	11.18 (33)	10.98 (58)
Deutsch (1,2)	.350	.961	.801	10.07 (80)	9.60 (5)	10.40 (58)	9.86 (37)	10.12 (180)
Gordon*	A#	A	.560	9.44 (9)	9.57 (7)	10.30 (87)	10.04 (56)	10.13 (159)
Gray (1-3)	A	A	.672##	9.00 (2)	9.11 (17)	9.67 (6)	8.79 (34)	8.98 (61)
Karnes (1-7)	.134	--	--	--	--	10.56 (22)	10.01 (75)	10.13 (97)
Levenstein*	A	A	.344##	10.93 (15)	10.91 (23)	10.49 (77)	10.57 (103)	10.61 (218)
Miller (1-5)	A	A	.317	10.67 (15)	10.29 (17)	10.88 (97)	10.89 (98)	10.82 (227)
Palmer	.258	A	.546##	11.42 (19)	11.27 (48)	11.47 (58)	11.13 (174)	11.23 (299)
Weikart	No Dropouts	--	.841	--	9.38 (65)	--	9.47 (58)	9.42 (123)

#"A" implies the F for this test was less than 1.

##Control group has higher mother's educational level.

* Gordon: Alachua County School District only. Levenstein: excluding Group 19.

Table A-16

Attrition on the School Record - Comparison of Hollingshead ISP Scores*

	Significance Levels			Mean ISP score				
				Control		Program		
Project (Groups)	Attrition Main Effect	Interaction	Final Samp. P-C Compar.	Dropout	Final Sample	Dropout	Final Sample	Overall Mean
Beller (1,2)	No Data on Dropouts		.753 #	--	56.67 (18)	--	57.95 (19)	57.32 (37)
Deutsch (1,2)	.306	.508	*t* not computed	65.88 (76)	69.00 (1)	66.74 (54)	64.52 (31)	65.93 (162)
Gordon **	A ##	No Data on Control Dropouts	--	--	--	63.70 (80)	64.22 (31)	63.84 (111)
Gray (1-3)	.185	.126	.615 #	73.00 (2)	69.84 (19)	63.60 (5)	70.52 (31)	69.77 (57)
Karnes (1-7)	.695	--	--	--	--	65.95 (22)	65.08 (71)	65.30 (93)
Levenstein **	.023	A	.784 #	64.27 (15)	62.83 (23)	66.30 (77)	63.48 (102)	64.47 (217)
Miller (1-5)	.800	No Data on Controls	--	--	--	62.92 (78)	63.34 (92)	63.15 (170)
Palmer	.048	A	.565 #	55.05 (19)	58.04 (48)	55.89 (57)	59.30 (171)	58.17 (295)
Weikart	No Dropouts	--	.376	--	69.09 (65)	--	67.81 (58)	67.92 (123)

---

* The Hollingshead Index of Social Position (ISP) ranges from 11 to 77, with "11" representing the highest social class and "77" representing the lowest social class.

# Control group has lower ISP score, that is higher SES.

## "A" implies the *F* for this test was less than 1.

** Gordon: Alachua County School District only. Levenstein: excluding Group 19.

Table A-17

Attrition on the School Record - Comparison of Pretest Stanford-Binet IQ Scores

Project (Groups)	Significance Levels			Mean IQ Scores				
				Control		Program		
	Attrition Main Effect	Interaction	Final Samp. P-C Compar.	Dropout	Final Sample	Dropout	Final Sample	Overall Mean
Beller (1,2)	A#	.078	.477 ##	87.05 (20)	93.70 (33)	93.57 (21)	91.33 (36)	91.69 (110)
Deutsch (1,2)	.483	.235	.207	91.57 (120)	88.33 (9)	91.97 (239)	94.49 (35)	91.99 (403)
Gordon**	No Pretest IQ	--	--	--	--	--	--	--
Gray (1-3)	A	A	.885	78.50 (2)	88.67 (15)	89.88 (8)	89.33 (36)	88.89 (61)
Karnes (1-7)	A	--	--	--	--	94.71 (24)	94.76 (78)	94.74 (102)
Levenstein*	A	A	.769	87.20 (5)	83.59 (22)	82.88 (75)	84.40 (101)	83.82 (203)
Miller (1-5)	No Pretest IQ	--	--	--	--	--	--	--
Palmer	.083	.079	<.001	86.50 (18)	84.84 (45)	89.86 (29)	95.87 (91)	91.28 (183)
Weikart	No Dropouts		.373		78.54 (65)		79.57 (58)	79.03 (123)

"A" implies the F for this test was less than 1.

Control group has higher pretest IQ score.

* PPVT pretest IQ. Analyses exclude Group 19.

** When Stanford Binet score at 3 years old (posttest IQ) is used, the F's for attrition main effect and interactions are both less than 1. Gordon: Alachua County School District only.

Table A-13

Attrition on the WISC-R - Comparison of Mother's Education

	Significance Levels			Mean Grades Completed				
				Control		Program		
Project (Groups)	Attrition Main Effect	Interaction	Final Samp. P-C Compar.	Dropout	Final Sample	Dropout	Final Sample	Overall Mean
Beller (1,2)	No WISC data has been received		--	--	--	--	--	--
Deutsch (1,2)	.842	.448	.706[##]	10.00 (69)	10.25 (16)	10.37 (46)	10.02 (49)	10.12 (180)
Gordon	.237	No data for Control Dropouts	.713	--	9.73 (11)	10.35 (85)	9.97 (58)	10.16 (154)
Gray (1-3)	A[#]	A	.814	10.00 (3)	8.94 (18)	8.75 (8)	8.97 (32)	8.98 (61)
Karnes (1-7)	.390	--	--	--	--	10.36 (22)	10.07 (75)	10.13 (97)
Levenstein	A	A	.573	10.55 (53)	10.12 (25)	10.62 (120)	10.40 (50)	10.51 (248)
Miller (1-5)	A	A	.171	10.67 (15)	10.29 (17)	10.89 (95)	10.87 (100)	10.82 (227)
Palmer	A	A	.765[##]	11.33 (39)	11.29 (28)	11.25 (129)	11.17 (103)	11.23 (299)
Weikart	.161	A	.777	10.11 (9)	9.27 (56)	10.50 (4)	9.39 (54)	9.42 (123)

[#] "A" implies the F for this test was less than 1.

[##] Control group has higher level of mother's education.

Table A-19

Attrition on the WISC-R - Comparison of Hollingshead ISP Scores*

	Significance Levels			Mean ISP Scores				
				Control		Program		
Project (Groups)	Attrition Main Effect	Interaction	Final Samp. P-C Compar.	Dropout	Final Sample	Dropout	Final Sample	Overall Mean
Beller (1,2)	No WISC data has been received		--	--	--	--	--	--
Deutsch (1,2)	.735	.888	.928	65.94 (69)	65.75 (8)	66.24 (45)	65.57 (40)	65.93 (162)
Gordon	A #	--	--	--	--	63.49 (79)	64.72 (32)	63.84 (111)
Gray (1-3)	A	A	.989 ##	70.33 (3)	70.11 (18)	66.86 (7)	70.21 (29)	69.77 (57)
Karnes (1-7)	.758	--	--	--	--	65.77 (22)	65.14 (71)	65.30 (93)
Levenstein	A	.032	.011	64.36 (53)	67.80 (25)	65.21 (119)	63.50 (50)	64.94 (247)
Miller (1-5)	.514	No Data on Controls	--	--	--	62.76 (76)	63.66 (96)	63.26 (172)
Palmer	.154	A	.440 ##	57.03 (39)	57.43 (28)	57.30 (128)	59.92 (100)	58.17 (295)
Weikart	A	A	.312	67.11 (9)	69.42 (55)	67.50 (4)	67.83 (54)	68.48 (122)

---

* The Hollingshead Index of Social Position (ISP) ranges from 11 to 77, with "11" representing the highest social class and "77" representing the lowest social class.

# "A" implies the F for this test was less than 1.

## Control group has lower ISP, i.e. higher SES level

Table A-20

Attrition on the WISC-R - Comparison of Pretest-Stanford Binet IQ Scores

	Significance Levels			Mean IQ Scores				
				Control		Program		
Project (Groups)	Attrition Main Effect	Interaction	Final Samp. P-C Compar.	Dropout	Final Sample	Dropout	Final Sample	Overall Mean
Beller (1,2)	No WISC-R data received		--	--	--	--	--	--
Deutsch (1,2)	.796	.873	.903	91.24 (112)	92.06 (17)	92.26 (227)	92.47 (47)	91.99 (403)
Gordon	No Pretest IQ		--	--	--	--	--	--
Gray (1-3)	A[#]	A	.761	87.33 (3)	87.50 (14)	91.40 (10)	88.85 (34)	88.89 (61)
Karnes (1-7)	.217	--	--	--	--	93.48 (23)	95.11 (79)	94.74 (102)
Levenstein[*]	.036	No Data on Follow-up Controls	--	83.83 (36)	--	82.74 (117)	86.32 (50)	83.82 (203)
Miller (1-5)	No Pretest IQ	--	--	--	--	--	--	--
Palmer	A	A	.001	85.22 (36)	85.44 (27)	93.40 (70)	95.84 (50)	91.28 (183)
Weikart	.174	A	.220	81.89 (9)	78.00 (56)	80.00 (4)	79.54 (54)	79.03 (123)

[#] "A" implies the F for this test was less than 1.

[*] PPVT pretest IQ.

APPENDIX B

Status Report by Instrument

The data collected from six instruments comprise the follow-up data base. Table B-1 shows the number of cases received from each project site by instrument as of the September 30, 1978 deadline.

Table B-1

Number of Cases Received by September 30, 1978 for Each Instrument by Project

Project Site	Parent Interviews	Youth Interviews	School Record Form	Ach't. Tests	Long School Record Form	IQ Scores*	Any Data**
Beller	108	109	107	102	102	--	121
Deutsch	81	107	73	63	--	107	119
Gordon	107	106	109	103	--	90	115
Gray	72	69	74	72	74	72	77
Karnes	165	156	153	143	105	112	168
Levenstein	98	75	75	115	--	76	188
Miller	141	141	141	134	--	141	141
Palmer	143	144	144	197	--	132	228
Weikart	106	104	104	96	--	110	123
Woolman	54	97	97	349	200	95	611
Zigler	--	185	185	185	--	185	185
Totals	1075	1293***	1851	1559	481	1120	2076

* Scores include WISC-R scores for Gordon, Gray, Karnes, Miller, Palmer, Woolman; WISC scores for Levenstein, Weikart; WAIS scores for Deutsch; and PPVT scores for Zigler.

** Any data is defined as number of cases for which data on at least one instrument were reported.

*** Total includes 185 Zigler cases who did not have Consortium Interview.

APPENDIX C

Response Rates and Final Dispositions

In order to calculate response rates for the follow-up sample, an attrition-disposition work sheet was sent to each project site. Eight sites (Beller, Gray, Karnes, Levenstein, Miller, Palmer, Weikart and Woolman) sent the data to Cornell in time to be included in this report. The final disposition of each ID case number in the eight sites was assigned to one of the following categories.

(C) Completed = Code 1

(R) Refused to give permission, refused to be interviewed = Code 6

(U) Located but unable to test because:
  Moved = Code 4
  Unable to test, keep appt., etc. = Code 5
  Terminated, unable to complete test = Code 7

(L) Lost, unable to trace on records = Code 3

(A) Attempted to locate and test: (U) + (L) = Codes 3 + 4 + 5 + 7

(OS) Out of sample
  Wave not scheduled to test at this follow-up = Code 2
  Dropped from sample at previous follow-up - Code 8
  No Data, unknown disposition = Code 9

Response rates are defined as the number of final dispositions in a given category divided by the number of cases in the total sample or subsample. The three response rates computed for completion and refusal categories are shown below:

Response Rate_1: Disposition divided by total possible cases

$$\text{e.g. } CR_1 = \frac{C}{C + R + A + OS}$$

Response Rate_2: Disposition divided by attempts (excludes out of samples)

$$\text{e.g. } CR_2 = \frac{C}{C + R + A}$$

Response Rate_3: Disposition divided by actual cases contacted (excludes lost)

$$\text{e.g. } CR_3 = \frac{C}{C + R + U}$$

The Response Rate Computations and Final Dispositions by Instrument for the eight project sites are shown in Table

As shown in Table C-1, out of the 1869 original subjects in the eight sites, 841 Parent Interviews were completed, resulting in a completion rate (1) of 44.9%. The rate of completion (2) based on the number of parents located and attempted to test is 67.0%. The completion rate (3) based on the parents actually contacted for this follow-up is 83.4%. The refusal rate (6) for the parents actually contacted is 3.2%. The refusal rate (5) based on the total number of parents attempted to locate is 2.6%. The refusal rate (4) based on all possible parents in the original sample is 1.7%. The total number of parent interviews attempted divided by the total number in the original population (7) yields a result of 67.0%. This result may be considered a sampling fraction or a measure of the effort to locate the parents which was dependent on time, money, luck, etc. In contrast, the sample not attempted (8) or out of sample remainder is 33.0%. As discussed in more detail in the technical supplement on attrition, the final samples are generally representative of the original samples in terms of differential rates of program and controls found, pretest Stanford-Binet IQ scores, Hollingshead ISP and mother's education.

The results are similar for the Youth Interview except that the refusal rate is lower. The refusal figure of 29 for the Youth Interview includes both youths who refused and parents who refused to have their child interviewed.

The refusal rates for the Youth Interview are 1.6% based on total original sample (4); 2.4% based on number located and attempted to test (5); and 2.9% based on subjects actually contacted (6). The completion rates for the Youth Interviews are: 45.6% based on total subjects in original sample (1); 69.2% based on located subjects (2); and 86.6% based on number of subjects actually contacted for this follow-up (3). The percentage attempted to 65.8% (7) and the out of sample percentage is 34.2% (8).

The Wechsler IQ response rates are similar to the Youth Interview figures. The School Record and Achievement have a much higher completion and attempted rate primarily because the Woolman site had access to all school records and the Weikart site sent data on all school recrords as of fourth grade.

In summary, the refusal percentages of 2-3% are acceptably low and the completion percentages appear satisfactory given the financial and time constraints.

Table C-1

Response Rate Computations and Final Dispositions for Eight Sites

Disposition	Rate Equation	Parent Interview		Youth Interview		School Record		Ach't. Test		WISC	
		N	%	N	%	N	%	N	%	N	%
(1) Completions	$CR_1 = \frac{C}{C + R + A + OS}$	841	44.9	853	45.6	1402	74.9	1159	61.9	738	39.4
(2) Completions	$CR_2 = \frac{C}{C + R + A}$	841	67.0	853	69.2	1402	82.9	1159	77.3	738	71.2
(3) Completions	$CR_3 = \frac{C}{C + R + U}$	841	83.4	853	86.6	1402	92.4	1159	88.7	738	87.3
(4) Refusals	$RR_1 = \frac{C}{C + R + A + OS}$	32	1.7	29	1.6	22	1.2	25	1.3	22	1.2
(5) Refusals	$RR_2 = \frac{C}{C + R + A}$	32	2.6	29	2.4	22	1.3	25	1.7	22	2.1
(6) Refusals	$RR_3 = \frac{C}{C + R + U}$	32	3.2	29	2.9	22	1.5	25	1.9	22	2.6
(7) Attempts	$AR = \frac{A + C + R}{C + R + A + OS}$	1255	67.0	1232	65.8	1691	90.3	1499	80.1	1036	55.3
(8) Not Attempted (out of sample)	$OS = \frac{OS}{C + R + A + OS}$	617	33.0	640	34.2	181	9.7	272	19.9	836	44.7

APPENDIX D

Special Analyses: Institute for Developmental Studies
(Drs. Cynthia and Martin Deutsch)

The Deutsch project was established in 1958 in public schools in lower Manhattan and East and Central Harlem. It evolved into a five year enrichment curriculum which ran from prekindergarten through third grade over about a 10 year period. Eight cohorts of children began the IDS program.

> From the original volunteers, children were randomly assigned to an experimental (group 1) and a "self-selected" control group (group 2)...who started school in regular public kindergarten. (Deutsch, et al., in Ryan, 1974)

In 1976, early data on over 1,000 children who participated in the IDS program (waves one through four) were sent to Cornell for Consortium analysis. These data included background characteristics and pretest and posttest IQ scores. Of this number, 504 children were in groups 1 and 2.

In 1976-78, follow-up data were collected on 119 children (groups 1-6, waves 1-4, and waves 5-8). The follow-up total includes 22 cases (treatment only) in waves 5-8 who were not on Consortium computer file previously. The analyses discussed in this report were limited to cohorts 1 through 4, group 1 (experimental) and group 2 (control). These groups and cohorts (N = 91 at follow-up, N = 504 at time of program) were randomly assigned and thus provide the most powerful evaluation of the IDS program.

Two analyses are presented in this section: school outcomes and Wechsler IQ scores.

School record information was received for 51 children (10.1% of the original 504 children). Data on special education assignment were available for 37 cases. One of the eight control children (12.5%) was assigned to a special education class, whereas none of the 29 treatment children was so assigned (Fisher exact p = .4324, two-tailed).

For grade retention, data were available for 37 children. For the treatment group, 23.3% (7 out of 30 children) were retained in grade one or more times compared to 42.9% of control children (3 out of 7 children). This difference results in a chi-square (without Yates correction) of 1.10, p = .2949, two-tailed.

The third dependent variable in the school outcomes analysis was underachievement. This was defined as assignment to special education classes and/or retained in grade and/or dropped out of school. Data were available for 43 of the children from the Deutsch project. This composite variable showed a marginally significant difference between treatment and control children. For the treatment group, 30.3% (10 out of 33 children) were "underachievers" compared to 60% of the control group (6 out of 10 children). This difference results in a chi-square of 2.90 and p = .0886, two-tailed. This result translates into an educationally significant difference in percent reduction terms.

When the underachiever percentages are compared, participation in the IDS program resulted in a 49.5% reduction in school failures ("underachievement").[1]

In summary, children who participated in the IDS program were more likely to meet the minimal requirements of their schools. As more data becomes available, we will be able to do further analyses.

Follow-up data on intelligence test scores were collected for 82 children (54 treatment and 28 controls). Most of these -- 61% -- were over 16 years old and received the Wechsler Adult Intelligence Scale (WAIS). The other children received the WISC-R. Separate analyses were planned for the young and the older children. However, because only three control children were given the WISC-R test, analyses were not done for the younger children.

A cautionary note must precede discussion of these results. As described in the attrition analyses, the children who were given the follow-up WAIS tests were significantly different from the original sample on the Stanford-Binet scores at 5 years old. The control children who were found had a lower mean IQ score at 5 than the controls who were not found; and the found treatment children had a higher IQ score at 5 years than the treatment children not found. An additional problem for the WAIS analysis was that sex was confounded with treatment: the treatment children were predominantly female.

---

[1] Percent reduction = (% control-% treatment)/% control.

The WAIS was administered to 26 treatment children and 24 control children. The mean age was 18 years 6 months. In general, the treatment children had higher mean scores than the control children but the results were not significant. As with older projects, the t-tests revealed no significant differences between treatment and control children on full scale, verbal or performance IQ scores (p = .08, .21, .10, two-tailed, respectively). None of the subtests showed significant differences.

APPENDIX E

Interview Coding and Analyses of Construct Validity

List of Interview Items and Coding of Dependent Variables

This appendix supplements the "Non-cognitive Outcomes: Attitudes and Values" section by providing a complete description of the coding of each variable. Variables are presented in the same order as in Table 16, and the same numbers and nomenclature are used.

Each variable is followed by a verbatim quote of the interview item used and by miscellaneous explanatory notes.

1. Mother's occupational aspirations for the child. "What kind of job would you like (child's name) to have later in life?"

Specific preferences were coded according to the Hollingshead Scale. An additional category was provided for responses which indicated that the decision was the child's and which listed no preference.

2. Percent no choice. See #1. This variable was the percentage of responses indicating that the child's vocation was the child's own decision.

3. (Algebraic) discrepancy scores. Simple difference between #1 and #2.

4. Absolute discrepancy scores. Absolute magnitude of the difference between #1 and #2.

5. Educational expectations. "How far do you plan to go in school?" This item was coded as an eight-point ordinal scale:

1 complete grammar school
2 some vocational high school
3 some regular high school
4 complete vocational high school
5 complete regular high school
6 some college
7 complete college
8 graduate or professional training
9 missing

6. Occupational aspirations. "What kind of job do you want to have as an adult?" This item was coded according to the seven-point Hollingshead Scale (Hollingshead, 1957).

7. Whether and where employed. "Do you do any kind of work for which you get paid?" (Within the last year.)

Coded according to a three-point ordinal scale:

0 no
1 yes, at home
2 yes, outside the home, or yes, both inside and outside the home

8. Amount of paid work. A composite of #7 and "How often do you work?"

Coded according to a four-point ordinal scale:

0 not at all
1 occasional, temporary, summer
2 part-time
3 full-time

9. Spare time: achievement. "What do you do in your spare time?" Up to three responses were coded, and multiple responses were summed. Responses were dichotomized as follows:

1 achievement-related
  - look for a job, working, babysit
  - hobbies, crafts, sewing, fix things

- do homework, study
- draw, writing stories (creative)
- practice musical instrument, singing
- sports activities (if participating)
- go to the library
- read, write letters

0 other
- play with friends, talk with friends, go places, get high
- watching sports activities
- other outdoor play
- go to social activities
- visiting friends, boy(girl) friends, relatives
- talk on phone
- play indoor games
- watch TV
- listen to music
- sleep, eat
- spend time by oneself
- housework, take care of children
- nothing
- no spare time

10; 11. Best things: achievement and worst things: achievement. "What is the best (worst) thing about school?" Up to three responses to each question were recorded and summed. Coding of the two variables was identical, except that the worst things variable was assigned a negative weight. Responses were coded dichotomously as follows:

1 achievement-related
- academic subjects, general academic activity
- non-academic subjects, including P.E., library
- unspecified learning (unless "easy work")[1]
- homework, exams[1]
- internships[1]
- daily attendance, promptness
- responsibilities

---

[1] As originally coded, these categories included responses which were inconsistent from the point of view of the present analysis. Therefore, the classification of responses in these categories as achievement-related or other was based on the original responses rather than the category *per se*. Accordingly, the placement of these categories in the above list is approximate.

0 other
- easy work
- field trips[1]
- social activities
- lunch
- going home, getting out, etc.
- fighting
- being out of the house
- teachers, other staff, how they treated me
- discipline, other students' behavior or attitudes
- unspecified activity, changing classes, etc.
- learning pace
- overcrowding
- nothing, everything

12. Proud: achievement. "Tell me something you've done to make you feel proud of yourself."

Coded dichotomously:

1 achievement-related
- school-related achievement, going to school
- job-related achievement, getting a job
- helping out at home
- sports achievement
- did well in some kind of competition; medal in Sunday School
- developed skills, joined group to better self or skill
- straightened oneself out
- doing better in everything

0 other
- found money
- good behavior[1]
- self-assertion[1]
- got married
- had a baby
- moved out on own
- altruistic acts
- interpersonal relations
- going to church
- nothing, everything

---

[1] As originally coded, these categories included responses which were inconsistent from the point of view of the present analysis. Therefore, the classification of responses in these categories as achievement-related or other was based on the original responses rather than the category *per se*. Accordingly, the placement of these categories in the above list is approximate.

13. Achievement-orientation ratings of reasons for admiring most admired adult. Answer to, "Why?", following "Of all the grown-ups you know personally, whom do you admire most? That is, who would you most like to be like in some way when you are older?"

Both polychotomous and dichotomous codings were employed. The former had nine categories:

- admires no adult
- because of richness, indulgence of subject, and other personality traits (not subsumed by the categories below)
- because of life style (e.g., has own apartment)
- because of occupational success
- because of education-related attributes
- because of skills other than education-related
- because teaches skills
- because of fame or respect from others
- because of unspecified achievement

The dichotomous code was created by collapsing the code above; responses were classified as achievement-related if they focused on education, knowledge, skills, occupational success, unspecified achievemant, or fame.

14. Self-evaluation of school performance. "How are you doing (did you do) in your schoolwork; that is, overall, not just in one subject? Is your schoolwork..." Following this, the interviewer read five alternatives which defined a five-point ordinal scale:

1 much better than the others in your classes
2 a little better than the others
3 about the same as others
4 a little worse than others
5 much worse than others

15. Self-evaluation of how well one gets along with the people with whom one lives. "Generally speaking, how do you get along with (people whom respondent has specified as those with whom he lives)? Would you say you get along..." Following this, the interviewer read five alternatives which defined a five-point ordinal scale:

1 very well
2 well
3 about average
4 not too well
5 badly

16. Participation in organized community activities. "Do you participate in any school, church, or community activities or belong to any groups or clubs like Scouts, sports, or the band?"

Responses were recorded as yes or no.

17. Frequency of participation in organized community activities. A composite of #16 and a subsequent question: "How often do they meet?"

Responses were coded into a six-point ordinal scale:

1 every day (5 or more days)
2 twice a week
3 once a week
4 twice a month
5 once a month
6 less than once a month

These values were transformed into a scale of times per month; category 1 was assigned a value of 20, and category 6 a value of 0.5. A value of zero was assigned if #16 was "no".

18. Whether one has "special friends" with whom one spends time. "Do you have any special friends that you spend time with?"

Responses were recorded as yes or no.

19. Frequency with which one gets together with "special friends." A composite of #18 and the subsequent question, "How often are you with them?"

Initial coding and subsequent tranformation of this item were identical to #17.

20; 21. Best things: sociability and worst things: sociability. As #10 and #11, except that the dichotomous coding was:

1 sociability-related
  - being with friends
  - interpersonal relationships
  - recess, free periods
  - after school activities
  - going home
  - meeting new people
  - lunch
  - general (unspecified)
  - being out of the house
  - playing games

0 other
  - academic and non-academic subjects
  - fighting
  - unspecified learning
  - teachers, other staff, the way they treated me
  - homework
  - field trips
  - internships
  - discipline
  - daily attendance
  - school activity, changing classes
  - pace of learning
  - overcrowding of facilities
  - responsibilities
  - nothing, everything
  - finishing, etc.

22. Spare time: sociability. As #9, but the dichotomous code was:

1 sociability-related
  - play with friends (unspecified), etc.
  - go to social activities

- visit friends, boy(girl) friend, relatives
- talk, talk on phone
- play indoor games

0 other

- sports activities
- watching sports activities
- other outdoor play (e.g., ride bike)
- go to library
- look for a job, work, babysit
- watch TV
- read, writing letters
- hobbies, etc.
- do homework, study
- draw, write (creative) stories
- listen to music
- eat
- sleep
- practice musical instrument, sing
- spend time by oneself
- housework, take care of children, etc.
- nothing
- no spare time

Analyses of Construct Validity

The dependent variables analyzed in this report vary in their face validity. In particular, the coding of several dichotomous achievement orientation variables involved considerable conceptual ambiguity. (See section on Non-cognitive Outcomes in text.) This appendix, therefore, presents a test of the construct validity of two of those dichotomous variables (the achievement-orientation ratings of the respondents' spare-time activities and reasons for being proud of themselves), as well as the subjects' self-evaluation of their school performance and the discrepancy between their vocational aspirations and their mothers' aspirations for them.

One approach to testing the validity of variables within a domain--e.g., the items which are classified here as reflective of achievement orientation--is to assess the degree of intercorrelation between them. While such an approach to the issue of validity has clear value in some types of psychometric work, it may be seriously misleading when applied to motivational data. McClelland (1975) has argued persuasively that if several behaviors are expressions of a single motivational construct, it will often not be the case that the various behaviors are highly correlated with each other. Rather, he maintains that it is often more appropriate to view such behaviors as alternative expressions of the motivational state, so that they need not occur together. Indeed, they may be negatively correlated, if only because the individual has only a finite amount of time to divide between the various alternative expressions. For example, an intensely competitive individual who is high on achievement motivation, as McClelland (1955) defined it, may single-mindedly devote him or herself to academic competition in the hope of gaining national professional reknown, while showing no interest in (and devoting no time to) other competitive endeavors. In the context of our data, a subject may be intensely motivated to do well in school and may accordingly use his sparetime as a chance to relax ("I go fishing with Grandaddy") rather than to join his peers in competitive athletics.

An alternative approach is to test some form of predictive validity. Since the focus of the present definition of achievement orientation is those aspects of the construct which "bear on school performance," a correlation between school performance or achievement and the achievement orientation variables would be evidence of construct validity.

As mentioned above, the Consortium data include three measures of actual school achievement: assignment to special education, retention in grade, and (in two projects) dropping out of school. Analyses indicate that treatment has sizable effects on the rate of assignment to special education and lesser effects on retention. For the present purposes, all three of these variables were pooled to form a single dichotomous variable labelled "achiever." Subjects were classified as achievers if they had never been retained or assigned to special education and had not dropped out of school, while all those who failed to meet one or more of these criteria were classified as non-achievers.

In the case of either variables which purport to reflect the subject's achievement orientation or variables (e.g., discrepancy scores and self-evaluation of school performance) which would be thought to be related to school performance for other reasons, a strong correlation with this achiever variable would be convincing evidence of construct validity. The absence of a strong correlation, however, would reveal nothing definitive, for such a result could reflect two factors other than low validity:

1. The true correlation between school performance and the variable in question could be attenuated by coarseness of measurement. The dichotomous "achiever" variable is a very crude index of school performance; likewise, a great deal of information is lost in collapsing the wealth of responses to an interview item into a dichotomy. While such coarseness of measurement on the part of the interview items would indicate low validity (just as a variable can not be both unreliable and valid), the coarseness of the school performance measure is irrelevant to the question of validity.

2. The size of the true correlation between real school performance and the non-cognitive variables considered here is

unknown. That is, even if the variables were perfectly valid and reliable and were highly refined (rather than crude dichotomies), the correlations might be small, for it might be that most of the variance in school performance in these samples is predicted by other variables--e.g., income, birth order, mothers' education, quality of the schools, and so on. The most appropriate test of validity would be the degree to which the observed correlations approach these unknown true correlations, not the degree to which the former approach 1.00.

Table E-1 presents the phi coefficients between the achiever variable and the respondents' reasons for being proud of themselves. Across all projects, the association is significant ($p < .02$, two-tailed), which demonstrates some validity of the item, as coded, as an index of achievement-orientation.

The pattern of age differences revealed by Table D-1 is striking: the association among the older projects is positive (mean phi = .26) and highly significant ($p = .0005$, two-tailed), while there is essentially no association in the younger projects. Indeed, two of the three non-significant (and very small) differences in the younger subsample went in the contrary direction (i.e., non-achievers scored higher on the interview item), and the interaction between age and the achiever variable as predictors of the interview item is significant at the .01 level (two-tailed).

Table D-2 presents means and t-tests comparing achievers' and non-achievers' self-evaluation of their school performance. Across all projects, the self-evaluations of the treatment groups are significantly higher ($p < .003$, two-tailed).

A skeptic might point to the fact that many of the differences in Tables E-1 and E-2 are non-significant, as well as to the negative direction of some of the non-significant differences, and argue that despite the two sources of attenuation described above, the construct validity of these two variables has not been convincingly demonstrated. However, the consistency of the age differences manifested by these two variables argues in favor of either of two alternative viewpoints. First, it is possible that the construct validity of the two variables increases with age. Second, these data might indicate developmental changes in the relationship between school performance and certain related attitudes in these samples. It should be noted in this regard that the Weikart project, which among the older projects showed the weakest relationships in both cases, also has the youngest subjects. At follow-up, Weikart's subjects ranged from 15 to 19 years old, while Beller's were 18 and Gray's, 19. These two possible interpretations are, of course, not mutually exclusive.

The interview item which posed the greatest ambiguity in coding was the achievement-orientation rating of spare-time activities. Table E-3 presents the differences between achiever and non-achiever on this item, as it was first coded, i.e., with participation in competitive athletics coded as achievement-related. It will be recalled that when the item was coded this way, control subjects were rated significantly more achievement-oriented. Table E-3 indicates that there is a trend of marginal significance (p = .097. two-tailed) favoring the achievers in the younger samples. In the older samples and overall, however, there was no significant difference. (Recall in this regard that the treatment-control difference on this item was due to the males in the younger projects.) Hence the data offer no firm evidence of construct validity.

It was noted earlier that the apparent negative treatment effect on this spare-time item, as originally coded, appeared to be due in large measure to the fact that competitive athletics were classified as achievement-related; when the responses were recoded so that athletics were omitted from the achievement-related category, no significant treatment-control difference remained. Accordingly, it was decided to test the construct validity of this second recoding as well. Table D-4 presents the results. It can be seen that when athletics were not classified as achievement-related, "achievers" were rated slightly more achievement-oriented when all projects were pooled ($p = .043$, two-tailed). This pattern is marginally significant in the younger projects ($p = .054$) but not in the older. Hence there is some slight evidence of the construct validity of this variable when the second recoding--omitting athletics from the achievement-related category--is employed.

However, it would be misleading to state simply that the second recoding (which yielded no treatment effect) was significantly associated with the achiever variable, while the first (which yielded a slight negative treatment effect) was not. One can calculate from Tables E-3 and E-4 that the mean correlation between the first recoding and the achiever variable was .051, while that yielded by the second recoding was .077. Neither of these figures is large enough to be persuasive, and the difference between them is even less so. One can only note with some caution that removing competitive athletics from the achievement-related category--and thus eliminating the slight negative treatment effect reported earlier--does, if anything, slightly enhance the none too substantial evidence of the variable's construct validity.

Table E-5 presents the mean discrepancy between mother's aspirations for their children and the children's own aspirations of both achievers and non-achievers in each project. It is clear that there is no significant overall association between the two measures, nor is there consistency across projects in the direction

In sum, the construct validity of four dependent variables was assessed. Some support was found for the validity of two of the measures: the "proud: achievement" variable and the respondents' self-evaluations of their school performance. Two recodings of the third variable--the "spare time: achievement" measure--were tested; in neither case was there persuasive evidence of construct validity, though the evidence was slightly more positive in the case of the recoding which excluded athletics from the achievement-related category. Finally, no evidence was found in support of the validity of the discrepancy score measure. These results must be interpreted in the light of several factors--explained above--which make the absence of a significant finding inconslusive.

Table E-1

Percent of School "Achievers" and "Non-achievers" Who Gave Achievement-Related Reasons for Being Proud of Themselves

Project (n)	Achievers	Non-Achievers	Difference	Φ	p (2-tailed)
Young Projects (ages 9-13)					
Gordon (67)	81.4%	87.5%	- 6.1%	-.079	(.52)[#]
Levenstein [‡‡]	--	--	--	--	--
Miller (125)	79.2	83.3	- 4.1	-.041	(.65)
Palmer (69) [‡]	78.0	68.4	9.6	.099	.41
Mean	79.5	79.7	- 0.2		(.876)
Old Projects (ages 15-19)					
Beller (89)	70.5%	53.3%	17.2%	.176	.097
Gray (50)	90.5	51.7	38.8	.410	.0037
Weikart (64)	86.4	70.0	16.4	.194	.120
Mean	82.5	58.3	24.2		.0005
All Projects	81.0%	69.0%	12.0%		.018

[‡] All male sample.

[‡‡] Data not available.

[#] Figures in parentheses are in the negative direction.

Table E-2

Mean of Children's Self-Evaluations of Their Overall School Performance Relative to Their Peers, "Achievers" vs. "Non-achievers"

1 = Much better than others in your classes;
3 = About the same as others;
5 = Much worse than others.

Project (n)	Achievers	Non-achievers	Difference	r_{pb}	p (2-tailed)
Young Projects (ages 9-13)					
Gordon (73)	2.84	2.69	.15*	-.097	(.414)#
Levenstein (75)	2.55	2.88	-.33	.150	.198
Miller (139)	2.85	2.81	.04	-.021	(.800)
Palmer ‡ (139)	2.62	2.74	-.12	.066	.441
Mean	2.72	2.78	-.06		.621
Old Projects (ages 15-19)					
Beller (90)	2.22	3.04	-.82	.463	.000005
Gray (50)	2.62	2.83	-.21	.197	.170
Weikart (69)	2.57	2.77	-.20	.121	.320
Mean	2.47	2.88	-.41		.00007
All Projects	2.61	2.82	-.21		.0028

‡ All male sample.

* A positive difference indicates that the self-evaluations of the control children were higher.

# Figures in parentheses are in the reverse direction.

Table E-3

Mean Achievement-orientation Ratings of Spare-time Activities, "Achievers" vs. "Non-achievers" (with competitive athletics coded as achievement-related)

Project (*n*)	Achievers	Non-achievers	Difference	r_{pb}	*p* (2-tailed)
Young Projects (ages 9-13)					
Gordon	.69	.58	.11	.079	.524
Levenstein ‡‡	--	--	--	--	--
Miller (139)	.66	.42	.24	.159	.063
Palmer ‡ (139)	1.05	1.00	.05	.032	.706
Mean	.80	.67	.13		.097
Old Projects (ages 15-19)					
Beller (89)	1.00	.86	.14	.087	.420
Gray (48)	.52	.67	-.15	-.111	(.452)#
Weikart (96)	.95	.87	.08	.057	.582
Mean	.82	.80	.02		.729
All Projects	.81	.73	.08		.157

‡ All male sample.

‡‡ Data not available.

# Figures in parentheses are in the reverse direction.

Table E-4

Mean Achievement-orientation Ratings of Spare-time Activities, "Achievers" vs. "Non-achievers" (omitting competitive athletics from the achievement-related category)

Project (n)	Achievers	Non-achievers	Difference	r_{pb}	p (2-tailed)
Young Projects (ages 9-13)					
Gordon (68)	.50	.42	.08	.058	.643
Levenstein ‡‡	--	--	--	--	--
Miller (139)	.37	.15	.22	.209	.016
Palmer ‡ (139)	.53	.47	.06	.041	.631
Mean	.47	.35	.12		.054
Old Projects (ages 15-19)					
Beller (89)	.76	.41	.35	.237	.025
Gray (48)	.29	.37	-.08	-.071	(.630)#
Weikart (96)	.31	.32	-.01	-.014	(.890)
Mean	.45	.37	.08		.352
All Projects	.46	.36	.10		.043

‡ All male sample.

‡‡ Data not available.

# Figures in parentheses are in the reverse direction.

Table E-5

Mean Discrepancy Between Mother's Occupational Aspirations for the Child and the Child's Own Aspirations, "Achievers" vs. "Non-achievers" (Based on Hollingshead Codes)*

Project (n)	Achievers	Non-achievers	Difference	r_{pb}	p (2-tailed)
Young Projects (ages 9-13)					
Gordon (47)	-.16	-1.13	.97	.245	(.101)[#]
Levenstein (27)	-.47	- .63	.16	.050	(.811)
Miller (111)	-.67	- .43	-.24	-.048	.624
Palmer ‡ (35)	-.93	- .13	-.80	-.204	.259
Mean	-.56	- .58	.02		(.897)
Old Projects (ages 15-19)					
Beller (60)	-.74	.21	-.95	-.269	.039
Gray (32)	-.21	-.39	.18	.044	(.818)
Weikart ‡‡	--	--	--	--	--
Mean	-.48	-.09	-.39		.197
All Projects	-.53	-.42	-.11		.529

* See Table 17 for Hollingshead code.

‡ All male sample.

‡‡ Data not usable.

# Figures in parentheses are in the reverse direction.

APPENDIX F

DEVELOPMENTAL CONTINUITY CONSORTIUM
PARENT FOLLOW-UP INTERVIEW

9/20/76)

CHILD NAME____________________ CHILD ID______________________ PARENT NAME_______________

[DO NOT READ ALTERNATIVE ANSWERS TO QUESTIONS UNLESS SPECIFICALLY INSTRUCTED TO DO SO.
RECORD 9 or 99 FOR NO ANSWER OR NOT APPLICABLE, 8 or 98 FOR DON'T KNOW]

1. First, I want to ask you some questions about your household. Could you list for me everyone who lives here, beginning with yourself and the other adults and going down to the youngest children? [RESPONDENT MAY LIST BY NUMBERS OR INITIALS, RATHER THAN NAMES IF DESIRED].

a. What relation is (person) to (sample child)?

b. Is that a man (boy) or woman (girl)? [ASK ONLY IF NOT OBVIOUS BY NAME]

c. How old is he/she? [AS OF LAST BIRTHDAY]
[IF UNDER 2 YEARS, RECORD AGE (UNDER 1 YR, RECORD 00) AND SKIP TO NEXT PERSON]

[IF 2 TO 6 YEARS, RECORD AGE AND ASK d:]

d. Is (person) in school (pre-school, nursery) part time or full time?
↑———— 0 No [RECORD AND SKIP TO NEXT PERSON]
1 Part time or 2 Full time [RECORD AND ASK e:]
e. What grade is he/she in? [RECORD 50 FOR pre-school, 51 FOR
↑———— Kindergarten, 01 FOR First grade, etc. AND SKIP TO NEXT PERSON]

[IF 6 TO 14 YEARS, RECORD AGE AND ASK d:]

d. Is (person) in school part time or full time?
0 No [ASK f:)
f. What grade did he/she complete? [RECORD GRADE IN YEARS AND SKIP
↑———— TO NEXT PERSON]
1 Part time or 2 Full time [RECORD AND ASK e:]
↑———— e. What grade is he/she in? [RECORD GRADE AND SKIP TO NEXT PERSON]

[IF OVER 14 YEARS, RECORD AGE AND ASK d:]

d. Is (person) in school part time or full time?

0 No [RECORD AND ASK f:]

f. What grade did he/she complete?
↓———— [IF LESS THAN 12 GRADES, RECORD AND SKIP TO h. below]
[IF 12 OR MORE GRADES COMPLETED ASK g:]

↓———— g. What degree or diploma did he/she receive?[SKIP TO h.]
1 Part time or 2 Full time [RECORD AND ASK e:]

e. What grade is he/she in? [RECORD AND ASK h:]

h. Does (person) have a part time or full time paid job?
0 No [IF UNEMPLOYED OR RETIRED, ASK:]

Has (person) ever had a paid job? [IF NO, ASK:] What is he doing now?
[IF YES, ASK i. BELOW]
1 Part time or 2 Full time [RECORD AND ASK i:]

i. What does (did) he/she do on their job? [PROBE FOR SPECIFIC HOLLINGSHEAD CODABLE JOB TITLE AND DESCRIPTION.

	Project ID	Child ID	File	Card	
____________Child ID	☐☐	☐☐☐☐	1	1	1-8

1. Household	Person 1 Name or # (Respondent) ________	Person 2 ________	Person 3 ________	Person 4 ________	Person 5 ________	Person 6 ________
a. Relation to Sample child: 01 Father 02 Mother 03 Guardian/Step-Parent 04 Grandparent 05 Uncle/Aunt 06 Gr. Unc/Aunt 07 Sibling 08 Hf/Step Sib 09 Cousin 10 No relation 11 Sample child 12 Other (specify)	☐☐ ________	☐☐ ________	☐☐ ________	☐☐ ________	☐☐ ________	☐☐ 9-20 ________
b. Sex 0 Male 1 Female	☐	☐	☐	☐	☐	☐ 21-26
c. Age in years	☐☐	☐☐	☐☐	☐☐	☐☐	☐☐ 27-38
d. In School 0 No 1 Part-time 2 Full-time	☐	☐	☐	☐	☐	☐ 39-44
e. Present Grade or f. Grade Completed (in years)	☐☐	☐☐	☐☐	☐☐	☐☐	☐☐ 45-46
g. Degree 0 No 1 Grad prof 2 BA 3 Some college 4 H.S. 5 Other (Specify degree)	☐ ________	☐ ________	☐ ________	☐ ________	☐ ________	☐ 57-62 ________
h. Working 0 No 1 Part-time 2 Full-time	☐	☐	☐	☐	☐	☐ 63-68
i. Kind of Job (if working) or other Activity	________ ________ ________	________ ________ ________	________ ________ ________	________ ________ ________	________ ________ ________	________ ________ ________
FOR OFFICE USE	☐	☐	☐	☐	☐	☐ 69-74

__________ Child ID

| * | * | * | * | * | * | * | 2 | 1-8

1. Household	Person 7 Name or # (Respondent)	Person 8	Person 9	Person 10	Person 11	Person 12
	______	______	______	______	______	______
a. Relation to Sample child: 01 Father 02 Mother 03 Guardian/Step-Parent 04 Grandparent 05 Uncle/Aunt 06 Gr. Unc/Aunt 07 Sibling 08 Hf/Step Sib 09 Cousin 10 No relation 11 Sample child	☐☐	☐☐	☐☐	☐☐	☐☐	☐☐ 9-20
12 Other (specify)	______	______	______	______	______	______
b. Sex 0 Male 1 Female	☐	☐	☐	☐	☐	☐ 21-26
c. Age in years	☐☐	☐☐	☐☐	☐☐	☐☐	☐☐ 27-38
d. In School 0 No 1 Part-time 2 Full-time	☐	☐	☐	☐	☐	☐ 39-44
e. Present Grade or f. Grade Completed (in years)	☐☐	☐☐	☐☐	☐☐	☐☐	☐☐ 45-46
g. Degree 0 No 1 Grad prof 2 BA 3 Some college 4 H.S. 5 Other (Specify degree)	☐ ______	☐ ______	☐ ______	☐ ______	☐ ______	☐ 57-62 ______
h. Working 0 No 1 Part-time 2 Full-time	☐	☐	☐	☐	☐	☐ 63-68
i. Kind of Job (if working) or other Activity	______ ______ ______ ______ ______	______ ______ ______ ______ ______	______ ______ ______ ______ ______	______ ______ ______ ______ ______	______ ______ ______ ______ ______	______ ______ ______ ______ ______
FOR OFFICE USE	☐	☐	☐	☐	☐	☐ 69-74

__________ Child ID

| * | * | * | * | * | * | * | 3 | 1-8

1. Household	Person 13 Name or # (Respondent)	Person 14	Person 15	Person 16	Person 17	Person 18	
	____	____	____	____	____	____	
a. Relation to Sample child: 01 Father 02 Mother 03 Guardian/Step-Parent 04 Grandparent 05 Uncle/Aunt 06 Gr. Unc/Aunt 07 Sibling 08 Hf/Step Sib 09 Cousin 10 No relation 11 Sample child	☐☐	☐☐	☐☐	☐☐	☐☐	☐☐	9-20
12 Other (specify)	____	____	____	____	____	____	
b. Sex 0 Male 1 Female	☐	☐	☐	☐	☐	☐	21-26
c. Age in years	☐☐	☐☐	☐☐	☐☐	☐☐	☐☐	27-38
d. In School 0 No 1 Part-time 2 Full-time	☐	☐	☐	☐	☐	☐	39-44
e. Present Grade or f. Grade Completed (in years)	☐☐	☐☐	☐☐	☐☐	☐☐	☐☐	45-46
g. Degree 0 No 4 H.S. 1 Grad prof 5 Other (Specify degree) 2 BA 3 Some (college)	☐ ____	☐ ____	☐ ____	☐ ____	☐ ____	☐ ____	57-62
h. Working 0 No 1 Part-time 2 Full-time	☐	☐	☐	☐	☐	☐	63-68
i. Kind of Job (if working) or other Activity							
FOR OFFICE USE	☐	☐	☐	☐	☐	☐	69-74

_____________ Child ID | * | * | * | * | * | * | * | 4 | 1-8

2. Does (child) have any other ~~full~~ brothers or sisters who are not living at home?
0 No
1 Yes [IF YES, ASK b-e] ☐ 9

b. How many brothers? ☐☐ 10-11

c. What are their ages?
[RECORD BEGINNING IN LEFTMOST BOX] ☐☐ ☐☐ ☐☐ ☐☐ 12-19

d. How many sisters? ☐☐ 20-21

e. What are their ages? ☐☐ ☐☐ ☐☐ ☐☐ 22-29

3.a Have there been any male adults who you would say have had a particularly important effect on (child) over the years?
0 No
1 Yes [IF YES, ASK:] ☐ 30

b. Who was that? [PROBE FOR RELATIONSHIP: IF PERSON NAMED IS OUTSIDE THE HOUSEHOLD, PROBE FOR MALE ADULT IN THE HOUSEHOLD] ☐☐ 31-32 ☐☐ 33-34

00 Natural father
01 Step-father
02 Grandfather
03 Uncle
04 Brother (full, half, or step)
05 Cousin
06 Male friend of mother
07 Friend of other family member
08 Social service agency representative (i.e. social worker, Big Brother)
09 Teacher, coach
10 Clergy
11 Other (specify)_____________

c. [WAS THIS PERSON NAMED AS HOUSEHOLD MEMBER IN QUESTION 1]
0 No
1 Yes ☐ 35

[IF NOT OBVIOUS, ASK:]

4. Do you live in a house or apartment?
1 House
2 Apartment
3 Mobile home
4 Duplex (2 family house with separate entrances) ☐ 36

[IF NOT OBVIOUS, ASK:]

5.a. Is it in a public housing project?
0 No
1 Yes [IF YES, ASK:] ☐ 37

b. What is the name of it?
PROJECT_____________________

6. Do you own or rent your house/apartment?
1 Own
2 Rent ☐ 38

7. Do you share your house/apartment with any other family?
0 No
1 Yes ☐ 39

8. How many rooms are there, not counting bathrooms, in your house (or part of the house you live in)? [IF RESPONSE INCLUDES HALF A ROOM, DISREGARD HALF]
☐☐ 40-41

_______________ Child ID

9. About how many houses or apartments have you lived in since (child) was five years old? ☐☐ 42-43

10. About how many schools has (child) been in since he/she started school? ☐☐ 44-45

[IF RESPONDENT NAMED A MALE ADULT IN QUESTION 1, ASK QUESTION 11, OTHERWISE SKIP TO QUESTION 12]

11. During the last year, would you say (male adult's name) has been employed?
[READ ALTERNATIVES] ☐ 46
1 not at all
2 1/4 of the year (3 months)
3 1/2 of the year (6 months)
4 3/4 of the year (9 months)
5 all of the year

12. During the last year, would you say you have been employed?
[READ ALTERNATIVES] ☐ 47
1 Not at all
2 1/4 of the year (3 months)
3 1/2 of the year (6 months)
4 3/4 of the year (9 months)
5 all of the year

Fine, that takes care of that part, now I would like to ask you some questions about (child). Some of these will be things that you may not have thought about before, so take as much time as you need to answer.

13. How far do you hope (child) will go in school?
[PROBE FOR AMOUNT] ☐☐ 48-49

[IF CHILD HAS DROPPED OUT OR GRADUATED, ASK:]

How far did you hope he/she would go?

01 Complete grammar school
02 some vocational high school
03 complete vocational high school
04 some regular high school
05 complete regular high school
06 some college
07 complete college
08 graduate or professional school
09 Other (specify)_______________
10-19 As far as he/she want to go (PROBE FOR AMOUNT: RECORD
  10 if amount not specified
  11 for grammar school
  12 some vocational school
  13 complete vocational school, etc.)

_______________ Child ID

[IF CHILD HAS DROPPED OUT OF SCHOOL, SKIP TO #15]
[IF CHILD HAS GRADUATED FROM HIGH SCHOOL, SKIP TO #14.b.]
[IF CHILD IS STILL IN SCHOOL, ASK # 14.a.]

14.a. Now considering how things are going in your family, how much schooling do you think (child) will be able to have? [RECORD RESPONSE AND PROBE FOR AMOUNT, THEN SKIP TO #15]

☐☐ 50-51

01 complete grammar school
02 some vocational high school
03 complete vocational high school
04 some regular high school
05 complete regular high school
06 some college
07 complete college
08 graduate or professional school
09 Other (specify)______________
____
10-19 As far as he/she wants to go [PROBE FOR AMOUNT--RECORD 10 if amount not specified, 11 for grammar school, etc.]

b. Do you think (child) will go any further in school?

☐ 52

0 No 1 Yes

[IF YES, ASK:]
Considering how things are going in your family, how much schooling do you think (child) will be able to have? [RECORD AND CONTINUE WITH #.15.]

06 some college
07 complete college
08 graduate or professional school
09 Other (specify)__________________
____
10-19 As far as he/she wants to go [PROBE FOR AMOUNT --RECORD 10 if no amount specified, 16 for some college, etc.]

☐☐ 53-54

15. How do you feel about how child has done (did) in school? [RECORD VERBATIM]

____
____
____
____

OFFICE CODE 1

☐☐ 55-56

Overall how satisfied are you with how (child) has done (did) in school? Would you say you are . . . [READ ALTERNATIVES]

1 Very satisfied
2 Somewhat satisfied
3 Somewhat dissatisfied
4 Very disassatisfied

☐ 57

16. What kind of job would you like (child) to have later in life? [PROBE FOR SPECIFIC JOB TITLE]

____
____
____

OFFICE CODE 2

☐ 58

______________ Child ID

17.a. Has (child) been sick a lot or had nervous problems?

0 No
1 Yes [IF YES, ASK:]

☐ 59

b. What was the trouble? ______________________________
____
____

OFFICE CODE 3
☐☐ 60-61

18.a. Has (child) ever spent more than a week in the hospital?

0 No
1 Yes [IF YES, ASK:]

☐ 62

b. Why was he/she in the hospital? ______________________
____

OFFICE CODE 3
☐☐ 63-64

c. How old was he/she then?

☐☐ 65-66

19.a. Now I want to talk about educational experiences some children have. Did (child) attend Head Start?'

0 No
1 Yes

☐ 67

[IF YES, ASK:]
b. For how long? [RECORD IN MONTHS]

☐☐ 68-69

20.a. Did (child) attend kindergarten?

0 No
1 Yes

☐ 70

[IF YES, ASK:]
b. For how long? [RECORD IN MONTHS]

☐☐ 71-72

☐* ☐* ☐* ☐* ☐* ☐* ☐* ☐5 1-8

________________ Child ID

a. ☐ 9

21.a. Did (child) attend Project Follow-Through?

0 No
1 Yes
[IF YES, ASK:]

b. ☐ 10
☐ 11
☐ 12
☐ 13

b. In what grades? [RECORD IN TOPMOST BOX AT b.]
0 Kindergarten, 1 1st grade, 2 2nd, 3 3rd grade

22.a. Did (child) participate in any unusual educational programs in his/her school like speech therapy, special education, etc. [INSERT LOCAL EXAMPLES]

☐ 14

0 No
1 Yes
[IF YES, ASK:]

b. Would you describe it for me?
[RECORD FIRST FOUR PROGRAMS MENTIONED]
0 Speech therapy
1 Special education
2 Vocational training
3 Advanced placement r programs for gifted
4 Remedial reading, math or other subject
5 Other (specify)________________________

____

____

Program	# of years	OFFICE CODE 4 Grades	
☐	☐☐	☐☐	15-19
☐	☐☐	☐☐	20-24
☐	☐☐	☐☐	25-29
☐	☐☐	☐☐	30-34

c. How many years was he/she in the program?

d. What grade was (child) in during the program? [EXPLAIN GRADE LEVEL(OR LEVELS) IN WHICH CHILD RECEIVED PROGRAM]________________

____

[OPEN ENVELOPE TO DETERMINE IF CHILD IS IN EXPERIMENTAL OR CONTROL GROUPS. IF EXPERIMENTAL, CONTINUE WITH QUESTIONS REGARDING INTERVENTION, IF CONTROL SAY:]

Those are all of the questions that I have for you. Thank you very much for your help. Do you have any questions that you would like to ask me about this study?

[DO NOT ASK NEXT QUESTIONS: COMPLETE AFTER END OF INTERVIEW]

23. [IF THE PARENT LIVES IN A PUBLIC HOUSING PROJECT, CALL HOUSING AUTHORITY TO DETERMINE IF IT IS:]

☐ 35

1 Low income
2 Moderate income
3 Mixed low and moderate income

24. [ETHNIC ORIGIN OF RESPONDENT]

0 Caucasian
1 Black
2 Oriental
3 Puerto Rican
4 Cuban
5 Other (specify)________________

☐ 36

[COMPLETE INTERVIEWER INFORMATION ON PAGE 13, LAST PAGE]

____________ Child ID

Delivery System

1 Center
2 Home
3 Combination

☐ 37

Now I would like to ask some specific questions about the program your child was in. These questions are just to help us evaluate that program; it doesn't mean that any new programs like it are being planned by us.

I'll be asking you to think back about the ____________ project that (child) was in several years ago, about what it was like, what it did for your child, things you liked about it and so on. I know that it has been a long time, but please try to remember. Take a few minutes to think about where you lived then, who was in the family--to help you remember back when (child) was that age.

25. What were some of the things you liked about the program? [RECORD VERBATIM]

OFFICE CODE 5
☐☐ 38-39
☐☐ 40-41
☐☐ 42-43
☐☐ 44-45

26. What did you like best? [RECORD VERBATIM]

27. What were some of the things you did not like about it? [RECORD VERBATIM]

OFFICE CODE 5
☐☐ 46-47
☐☐ 48-49
☐☐ 50-51
☐☐ 52-53
☐☐ 54-55
☐☐ 56-57

28. How would you have changed it to make it better? [RECORD VERBATIM]

________________ Child ID

29. Some educational programs for pre-school children are set up in centers or schools. Others are set up to have a home visitor bring educational activities to families in their own home.

[FOR COMBINATION HOME AND CENTER-BASED PROGRAMS, SKIP TO 30]
[FOR CENTER OR HOME-BASED PROGRAMS ONLY, ASK:]

a. Did you like having the program in your home/in a center?

0 No
1 Yes

☐ 58

Why or why not? [RECORD VERBATIM] ________________

OFFICE CODE 6
☐☐ 59-60
☐☐ 61-62
☐☐ 63-64

b. Would you have preferred to have had it in a center/your home?

0 No
1 Yes

☐ 65

c. Why or why not? [RECORD VERBATIM SKIP TO 31] ________________

OFFICE CODE 6
☐☐ 66-67
☐☐ 68-69
☐☐ 70-71

30.a. Did you like having the program in both a center and in your home?

0 No
1 Yes

☐ 72

b. Why or why not? [RECORD VERBATIM] ________________

OFFICE CODE 6
☐☐ 73-74
☐☐ 75-76
☐☐ 77-78

c. Did you prefer either the home or center part of the program?

0 No preference
1 Home
2 Center

*	*	*	*	*	*	*	6	1-8

☐ 9

[IF STATE A PREFERENCE, ASK:]
Why was that? [RECORD VERBATIM] ________________

OFFICE CODE 6
☐☐ 10-11
☐☐ 12-13
☐☐ 14-15

_______________ Child ID

31. Generally speaking, was the __________ program a good thing for your child?

☐ 16

0 No
1 Yes
8 Don't know, no answer, don't remember

32.a. Programs for young children are set up to provide different kinds of help to families. I have some specific questions about help you, your child, or your family may have received thru this __________ program. It may be hard to remember so take your time and think back to when (child) was that age.

Here is a list of services you may have received thru this __________ program. Look at each one and tell me which of these services or help you, your child or someone in your family received through the program.

[HAND RESPONDENT CARD WITH LIST OF SERVICES, AND ALLOW ADEQUATE TIME TO READ THE LIST.
IF RESPONDENT HAS QUESTIONS ABOUT THE SERVICE, EXPLAIN BY GIVING FURTHER EXAMPLES.]

00 No services received
01 Medical check-ups or tests?
02 Medical treatment like shots, medicine, physical therapy?
03 Help to find about ADC payments, food stamps, employment service, or the name of a dentist or doctor?
04 Rides or car fare to the doctor, for field trips to the program, for social services?
05 Help with personal or marital problems, education or job opportunities?
06 Classes for parents on home management, budgeting, taxes?
07 Parties, picnics, field trips for parents?
08 Information about how children grow and learn?
09 Help in getting along better with your child?
10 Help in learning how to teach your child things yourself?
11 Free time so that you could go to school, to work, or do shopping
98 Don't know
99 Not applicable

☐☐ 17-18
☐☐ 19-20
☐☐ 21-22
☐☐ 23-24
☐☐ 25-26
☐☐ 27-28
☐☐ 29-3
☐☐ 31-32
☐☐ 33-3
☐☐ 35-36
☐☐ 37-38

b. Now of the services you did receive through the program, which ones made a difference or were important to you?

[RECORD USING SAME CODE AS 32.a.]

[COMMENTS] ______________________________
____
____
____
____

b. ☐☐ 39-40
☐☐ 41-4
☐☐ 43-44
☐☐ 45-4
☐☐ 47-48
☐☐ 49-
☐☐ 51-52
☐☐ 53-54
☐☐ 55-56
☐☐ 57-58
☐☐ 59-60

That's all the questions I have for you. Thank you very much for you cooperation. Are there any questions you would like to ask me about the interview?

________________ Child ID

33. INTERVIEWER SIGNATURE ________________ INTERVIEWER ID ☐☐ 61-62

34. DATE OF INTERVIEW (Month, Day) ☐☐ ☐☐ 63-66

35. INTERVIEWER SEX 0 Male, 1 Female ☐ 67

36. INTERVIEWER ETHNIC ORIGIN ☐ 68
   0 Caucasian 3 Puerto Rican
   1 Black 4 Cuban
   2 Oriental 5 Other (Specify)

37. DURATION OF INTERVIEW (in minutes) ☐☐ 69-70

38. PLACE OF INTERVIEW ☐ 71
   0 Home of respondent
   1 Center site
   2 Telephone
   3 Other (Specify) ________________

39. RATE THE DEGREE OF INVOLVEMENT OF RESPONDENT ☐ 72
   1 Highly involved and interested
   2
   3 Neither involved nor uninvolved
   4
   5 Bored, not interested

COMMENTS ON PROBLEMS WITH INTERVIEW, SPECIAL CIRCUMSTANCES ________________

____

____

____

____

____

(9/20/76)

DEVELOPMENTAL CONTINUITY

YOUTH FOLLOW-UP INTERVIEW

CHILD'S NAME ____

PROJECT ID ☐☐ 1-2

CHILD ID ☐☐☐☐ 3-6

FILE, CARD [2][1] 7-8

Introductory Statement

I am a representative of the (name of the university or project). This is an independent organization--we are not part of the public schools or the government.

Our work involves learning more about young people like yourself. In this interview, we will be asking a number of questions about the things you like to do both in and outside school, about your family, your friends, and your life in general.

It is important that you feel free to tell me exactly what you think, and not just what you feel I want to hear. Everything you tell me is completely confidential. Your name will be removed from the interview form, so that no one will connect your name with your answers. Don't feel that we can't talk about the questions. If there are questions that aren't clear to you or that you feel are too personal, please tell me.

---

1. a. Let's talk about school first. Are you still in school? ☐ 9

   0 No [IF NO, ASK QUESTION 1b]
   1 Yes [IF YES, CONTINUE WITH QUESTIONS 1c AND 2]

   b. What is the last grade you completed and got credit for? [SKIP TO QUESTION 3] ☐☐ 10-11

   c. What grade are you in? ☐☐ 12-13

2. How far do you plan to go in school? [RECORD RESPONSE, PROBE FOR AMOUNT] ☐☐ 14-15

   01 complete grammar school
   02 some vocational high school
   03 complete vocational high school
   04 some regular high school
   05 complete regular high school
   06 some college
   07 complete college
   08 graduate or professional training
   10-19 as far as I can go [PROBE FOR AMOUNT; RECORD 10 IF AMOUNT NOT SPECIFIED, 11 FOR GRAMMER SCHOOL, 12 SOME VOCATIONAL, ETC.]

Child ID ____________

OFFICE CODE 7

☐☐ 16-17
☐☐ 18-19
☐☐ 20-21
☐☐ 22-23
☐☐ 24-25
☐☐ 26-27

3. What is (was) the best thing about school? [PROBE FOR SUBJECTS, INDICATE BEST THING WITH AN *]

____
____
____

4. What is (was) the worst thing about school? [PROBE FOR SUBJECTS, USE AN * TO INDICATE WORST]

____
____

5. How are you doing (did you do) in your schoolwork; that is, overall, not just in one subject? Is your schoolwork...[READ ALTERNATIVES]

1. Much better than the others in your classes
2. A little better than the others
3. About the same as others
4. A little worse than others
5. Much worse than others

☐ 28

6. a. Do you participate in any school, church or community activities or belong to any groups or clubs like Scouts, sports or the band?

0 No
1 Yes [IF YES, ASK:]

☐ 29

b. What is it (are they) called? [IF NOT CLEAR FROM THE NAME, PROBE: What kind of club is that?]

01 Scouts
02 Religious group
03 Sports
04 Music
05 Art, Photography
06 Newspaper
07 Social club
08 Service club (Red Cross)
09 Other (Specify) ____________
10 YMCA, Boys/Girls Clubs

b. Type of Club	c. How often	
☐☐	☐	30-32
☐☐	☐	33-35

c. How often do they meet?

0 Every day (5 or more days)
1 Twice a week
2 Once a week
3 Twice a month
4 Once a month
5 Less than once a month

7. a. Do you have any special friends that you spend time with?

0 No
1 Yes [IF YES, ASK:]

☐ 36

b. How often are you with them?

☐ 37

1 Every day(5 or more)
2 Twice a week
3 Once a week
4 Twice a month
5 Once a month
6 Less than once a month

8. What do you do in your spare time? ____________________
____

OFFICE CODE 8

☐☐ 38-39
☐☐ 40-41
☐☐ 42-43

Child ID ______________

9. a. Do you have any books of your own? [OTHER THAN TEXTBOOKS] ☐ 44

0 No
1 Yes

b. How many hours a week do you spend reading books or magazines? [RECORD 01 FOR 1 HOUR OR LESS] ☐☐ 45-46

c. What kinds of books do you read? [RECORD FIRST FOUR TYPES MENTIONED AND SPECIFY NAME OR SUBJECT OF BOOK]

01 Comic books ______________
02 Adventure/mystery ______________
03 Western ______________
04 Love/Romance ______________
05 Fantasy ______________
06 Spiritualism/Astrology ______
07 Science Fiction ______________
08 Pornography ______________
09 Amusement/Humor ______________
10 Games, Sports, Automobiles ___
12 Family Life ______________
13 Self Improvement ______________
14 Recreation/Hobbies ________
15 Information (Encyc.) ______
16 Biography/Autobiography ___
____
17 Arts and Music ______________
18 Nature/Science ______________
19 Animal Stories ______________
20 Other ______________
21 Crime ______________
22 Religious (Bible) __________
23 Other Nonfiction (History)
24 Other Fiction ______________

☐☐ 47-48
☐☐ 49-50
☐☐ 51-52
☐☐ 53-54

10. a. Do you read magazines? ☐ 55

0 No
1 Yes [IF YES ASK:]

b. What kinds of magazines do you read? [RECORD FIRST FOUR TYPES]

01 News (Time/Newsweek)
02 Love/Romance
03 Sports
04 Rock Music (Rolling Stone)
05 Hobbies (Arts and Crafts)
06 Adventure
07 Pornography
08 Fashion/Teenage advice (Seventeen, Glamour)
09 Feature (Ebony, Family Circle)
10 Comic Books
11 Children's Magazines (Jack & Jill)
12 Humor Magazines (Mad)
13 School Magazines (Weekly Reader, Jr. Scholastic)
14 Other (Specify) __________
____
15 Playboy ______________
16 Religious ______________

☐☐ 56-57
☐☐ 58-59
☐☐ 60-61
☐☐ 62-63

11. a. Do you read a newspaper? [THIS INCLUDES ANY PART OF NEWSPAPER] ☐ 64

0 No
1 Yes [IF YES, ASK:]

1. How often do you read it? ☐ 65

1 Every day
2 A couple of times a week
3 A couple of times a month

Child ID ____________

12. a. Do you use the school or community library?

0 No ☐ 66
1 Yes [IF YES, ASK: Which one?]
 1 School
 2 Community
 3 Both ☐ 67

b. Do you have a community library card?

0 No
1 Yes [IF YES, ASK:] ☐ 68

c. How many times a month do you use it? ☐ 69
 0 Never
 1 Once
 2 Twice
 3 Three times
 4 Four times
 5 More than four times
 6 Six to 11 times per year
 7 One to 5 times per year

13. a. Do you watch television?

0 No
1 Yes [IF YES, ASK:] ☐ 70
 b. How much television would you say you watch?
 Do you watch: [READ FIRST PHRASE OF ALTERNATIVES]
 1 More than three hours a day (21 + hours a week)
 2 2-3 hours a day (14-20 hrs. a weeks)
 3 1-2 hours a day (7-13 hrs. a week) ☐ 71
 4 Less than an hour a day (less than 7 hrs. a week)

14. Do you have any room in your home where you may go whenever you want to be alone?

0 No
1 Yes [IF YES, ASK:]
 Which room(s) is that? ☐ 72
 1 Own bedroom (or shared bedroom)
 2 Den or living room (playroom)
 3 Dining room
 4 Kitchen
 5 Cellar or attic ☐ 73
 6 Bathroom
 7 Other (Specify) ____________________

15. a. Do you do any kind of work for which you get paid?
[AT HOME OR OUTSIDE THE HOME, WITHIN THE LAST YEAR]

0 No [SKIP TO 16] ☐ 74
1 Yes, at home
2 Yes, outside the home [IF YES, ASK 15b and 15c]
3 Yes, both inside and outside the home

b. What kind of work do you do? [RECORD SPECIFIC JOB DESCRIPTION]

____

OFFICE CODE 9
☐ 75

Child ID ________________

15. (con't)

c. How often do you work? ☐ 76
1 Full-time
2 Part-time
3 Occasional, temporary, summer

*	*	*	*	*	*	*	2	1-8

16. What kind of job do you want to have . . .
[IF CHILD IS IN 8th GRADE OR LESS] when you grow up?

____

[IF CHILD IS IN 9th GRADE OR HIGHER] as an adult?

____

17. What is the worst trouble you've ever been in? [RECORD VERBATIM: PROBE: Anything else?]

____

____

____

18. Tell me something you've done that made you feel proud of yourself. [RECORD VERBATIM] ____________________

____

____

OFFICE CODE 10
☐ 9

OFFICE CODE 11
☐☐ 10-11

OFFICE CODE 12
☐☐ 12-13

OK, now let's talk a little about your family.

19. Different families have different types of living arrangements. Some young people live with one parent, others live with two parents, or with grandparents or relative. [IF MORE THAN ONE CODE, RECORD 14 and SPECIFY]

a. Who do you live with now? ☐☐ 14-15

00 Both parents
01 Mother only
02 Father only
03 Part of time with mother and part with father
04 Mother and step-father or substitute
05 Father and step-mother or substitute
06 Grandmother
07 Grandfather
08 Grandparents
09 Other relatives
10 Siblings
11 Foster family
12 Friends
13 Institution (reform school, orphanage)
14 Other (Specify) (include guardian if he/she is not one of other codes)
____
15 Living alone (or with own children)
16 Living with husband/wife

Child ID ______________

20. Have you ever lived somewhere other than with your parents (OR PERSON NAMED IN #19) for more than 3 months at one time?

0 No
1 Yes [IF YES, ASK:] ☐ 16

b. Who did you live with? [RECORD CODE AS IN #19]
OTHER (Specify) ________________________ ☐☐ 17-18

21. a. Do you eat supper alone or with other people, most of the time?

1. Alone
2. With others [IF WITH OTHERS, ASK:] ☐ 19

b. Who do you eat with?
1 Parent(s)
2 Sibling(s)
3 Friend (s)
4 Parent(s) and sibling(s)
5 Parent(s), sibling(s) and friend (s)
6 Other (specify) ______________ ☐ 20

c. Do people generally talk during the meal?
0 No
1 Yes ☐ 21

[ASK QUESTIONS 22-26 FOR WHOMEVER THE CHILD HAS SAID HE LIVES WITH IN QUESTION #19]

[IF PARENT IS PRESENT, ASK THEM TO LEAVE FOR QUESTION 22]

22. Generally speaking, how do you get along with (persons named in Question 19)? Would you say you get along ☐ 22

1 Very well
2 Well
3 About average
4 Not too well
5 Badly

23. Of all the grown-ups you know personally, whom do you admire most? That is, who would you most like to be like in some way when you are older? [PROBE; IF NO ANSWER, ASK: Is there anyone you would like to be like?] [DETERMINE RELATIONSHIP OF PERSON TO SUBJECT, AND DETERMINE WHY THE SUBJECT RESPECTS THE PERSON]

OFFICE CODE 13
☐☐ 23-24

____

____

____

24. Where have you and your family gone together within the city or area within the last month?

0 Shopping and other errands
1 Entertainment or outings (movies, ball games, parks, picnics, restaurants)
2 Church
3 Visit friends or relatives
4 Cultural events or places (museums, concerts)
5 Other (specify) __________
____
6 Didn't go anywhere with family

☐ 25
☐ 26
☐ 27

Child ID ____________________

25. a. Where have you and your family gone together out of the city or area in the last month? (RECORD PLACE AND PURPOSE)

00 Another area within the state
01 Another state in the same geographic area
02 Northeast U.S.
03 Southeast U.S.
04 Southwest U.S.
05 Mid-west U.S.
06 Northwest U.S.
07 Far west U.S.
08 Out of country
09 Not gone anywhere out of city or area

☐☐ 28-29

b. [PURPOSE OF TRIP. IF NOT MENTIONED, ASK] What did you do there?

01 Visit relative and friends
02 Vacation
03 Family emergency
04 Business trip (Shopping)
05 Household move
06 Other (specify) ____________________

☐☐ 30-31

26. How far away from home have you been without adults? [RECORD THE FARTHEST]

1 Out of the neighborhood
2 Out of the city
3 Out of the state
4 Out of country
5 Nowhere

☐ 32

[CHECK TO SEE IF CHILD IS EXPERIMENTAL OR CONTROL. IF CONTROL, THANK AND END INTERVIEW. IF EXPERIMENTAL, ASK:] Now to change the subject again, I want you to think back to when you were younger and were in the _________ program. Take your time.

27. Tell me what you remember about the ___________ program. [RECORD VERBATIM] [ALLOW ADEQUATE TIME. PROBE: Anything else?]

____

____

____

OFFICE CODE 6

☐☐ 33-34

☐☐ 35-36

Fine, that is the end of my questions. I appreciate your time and help. Are there any questions you would like to ask me?

INTERVIEWER SIGNATURE ____________________

DATE OF INTERVIEW (Mo., Day) ☐☐ ☐☐

DURATION OF INTERVIEW (in minutes) ☐☐

INTERVIEWER NUMBER ☐☐ 37-38

INTERVIEWER SEX 0 Male, 1 Female ☐ 39-43

INTERVIEWER ETHNIC ORIGIN ☐ 44-46

0 Caucasian
1 Black
2 Oriental
3 Puerto Rican
4 Cuban
5 Other (Specify)

____

1/28/77

SCHOOL AND ACHIEVEMENT TEST RECORD FORM

Child's name ______________________ Project ID ☐☐ (1-2) Child ID ☐☐☐☐ (3-6) File [4] (7)

Complete this section only once for each child.

1. What is the child's current status in school?
   - --if still in school, record in col. 8-9 the grade child is in, i.e. 07, 13 (for one year post-high school), etc. If ungraded, record 52. ☐☐ 8-9
   - --if graduated from high school and not presently in school, record 92 in col. 8-9.
   - --if dropped out, record 00 in col. 8-9; record grade dropped out in col. 10-11. ☐☐ 10-11

2. Was the child ever in special education of any type? ☐ 12
   - 1 No
   - 2 Special education (unspecified)
   - 3 EMR, TMR
   - 4 Learning disabled
   - 5 Emotionally disturbed
   - 6 Speech/hearing
   - 0 No information

3. Was the child ever retained or skipped a grade? ☐ 13
   - 1 No
   - 2 Retained once
   - 3 Retained twice or more
   - 4 Skipped once or more
   - 0 No information

Complete the remainder of the form for each of four years.

Report scores for each relevant year for your study. For example, studies with subjects who are currently high school age should record the latest year available and the 10th, 8th, and 4th grade scores. Studies with younger children should record the four latest years available, such as 7th, 6th, 5th, and 4th (see instructions).

List the names and identification information for each school the child attended when (s)he took the achievement tests to be recorded.

	Latest grade	Next highest grade	Next grade	Lowest grade	
4. Grade (ungraded = year e.g., 74)	☐☐	☐☐	☐☐	☐☐	14-21
5. School code or name	☐☐☐ ________	☐☐☐ ________	☐☐☐ ________	☐☐☐ ________	22-33
6. School district name or number	________	________	________	________	
				OFFICE USE ☐☐☐☐	34-37
7. School type 1 Public 2 Parochial 3 Private 0 No information	☐	☐	☐	☐	38-41
8. School location 1 Urban 2 Suburban 3 Small city 4 Rural 0 No information	☐	☐	☐	☐	42-45

CONSORTIUM ACHIEVEMENT TEST RECORD FORM

Child's name____

project ID	child ID	file #	grade	test code	month & year administered	battery or level	edition 19__	test form	grade norm	type norm
☐☐	☐☐☐☐	5	☐☐	☐☐	☐☐☐☐	☐☐	☐☐	☐	☐☐.☐	☐
1-2	3-6	7	8-9	10-11	12-15	16-17	18-19	20	21-23	24

name of subtest		raw score	percentile	stanine	grade=1 age=2 equiv.	grade or age equivalency score	standard score
	total test	☐☐☐ 25-27	☐☐ 28-29	☐ 30	☐ 31	☐☐.☐ 32-34	☐☐☐ 35-37
☐ 38	reading subtest	☐☐☐ 39-41	☐☐ 42-43	☐ 44	☐ 45	☐☐.☐ 46-48	☐☐☐ 49-51
☐ 52	reading subtest	☐☐☐ 53-55	☐☐ 56-57	☐ 58	☐ 59	☐☐.☐ 60-62	☐☐☐ 63-65
☐ 66	math subtest	☐☐☐ 67-69	☐☐ 70-71	☐ 72	☐ 73	☐☐.☐ 74-76	☐☐☐ 77-79

---

project ID	child ID	file #	grade	test code	month & year administered	battery or level	edition 19__	test form	grade norm	type norm
☐☐	☐☐☐☐	5	☐☐	☐☐	☐☐☐☐	☐☐	☐☐	☐	☐☐.☐	☐
1-2	3-6	7	8-9	10-11	12-15	16-17	18-19	20	21-23	24

name of subtest		raw score	percentile	stanine	grade=1 age=2 equiv.	grade or age equivalency score	standard score
	total test	☐☐☐ 25-27	☐☐ 28-29	☐ 30	☐ 31	☐☐.☐ 32-34	☐☐☐ 35-37
☐ 38	reading subtest	☐☐☐ 39-41	☐☐ 42-43	☐ 44	☐ 45	☐☐.☐ 46-48	☐☐☐ 49-51
☐ 52	reading subtest	☐☐☐ 53-55	☐☐ 56-57	☐ 58	☐ 59	☐☐.☐ 60-62	☐☐☐ 63-65
☐ 66	math subtest	☐☐☐ 67-69	☐☐ 70-71	☐ 72	☐ 73	☐☐.☐ 74-76	☐☐☐ 77-79

☆U.S. GOVERNMENT PRINTING OFFICE: 1979— 281-683/8032